FEDERAL RESERVE BANK OF ST. LOUIS

REVIEW

THIRD QUARTER 2020
VOLUME 102 | NUMBER 3

REVIEW

Volume 102 • Number 3

FEDERAL RESERVE BANK *of* ST. LOUIS

CENTRAL TO AMERICA'S ECONOMY®

Review

Review is published four times per year by the Research Division of the Federal Reserve Bank of St. Louis. Complimentary print subscriptions are available to U.S. addresses only. Full online access is available to all, free of charge.

Online Access to Current and Past Issues

The current issue and past issues dating back to 1967 may be accessed through our Research Division website: **http://research.stlouisfed.org/publications/review**. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* also are available to our readers on this website.

Review articles published before 1967 may be accessed through our digital archive, FRASER: **http://fraser.stlouisfed.org/publication/?pid=820**.

Review is indexed in Fed in Print, the catalog of Federal Reserve publications (**http://www.fedinprint.org/**), and in IDEAS/RePEc, the free online bibliography hosted by the Research Division (**http://ideas.repec.org/**).

Authorship and Disclaimer

The majority of research published in *Review* is authored by economists on staff at the Federal Reserve Bank of St. Louis. Visiting scholars and others affiliated with the St. Louis Fed or the Federal Reserve System occasionally provide content as well. *Review* does not accept unsolicited manuscripts for publication.

The views expressed in *Review* are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Subscriptions and Alerts

Single-copy subscriptions (U.S. addresses only) are available free of charge. Subscribe here: **https://research.stlouisfed.org/publications/review/subscribe/**.

Our monthly email newsletter keeps you informed when new issues of *Review*, *Economic Synopses*, *Regional Economist*, and other publications become available; it also alerts you to new or enhanced data and information services provided by the St. Louis Fed. Subscribe to the newsletter here: **http://research.stlouisfed.org/newsletter-subscribe.html**.

Economic Data

General economic data can be obtained through FRED®, our free database with over 760,000 national, international, and regional data series, including data for our own Eighth Federal Reserve District. You may access FRED through our website: **https://fred.stlouisfed.org**.

ISSN 0014-9187

A Short Tour of Global Risks

Carmen M. Reinhart

This article is based on the author's Homer Jones Memorial Lecture delivered at the Federal Reserve Bank of St. Louis, Wednesday, June 25, 2019. (JEL E51, F3, G15, G28)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2020, *102*(3), pp. 221-35.
https://doi.org/10.20955/r.102.221-35

It's a pleasure and an honor to deliver the 2019 Homer Jones Lecture. What I'd like to do is examine global risks and connect those risks to the literature and work that I've done. The tour begins with some of the risks. This is not meant to be encyclopedic; but I will try to be brief so we can cover a lot of ground—some of the risks in the advanced economies and then, truly global in nature, move on to risks in emerging markets. I would note that, in the past, interest in emerging markets was really limited to traders that bought emerging market bonds and occasionally equity. But while in the early 1980s emerging markets accounted for about a third of global gross domestic product (GDP), now they account for about two-thirds of global GDP. So, it's really not possible to talk about the global economy without a full, rounded view of advanced and emerging economies.

I'll first focus on what I see as more short-term concerns and then talk about a long-term issue that keeps cropping up on my radar screen—something that I've been working on for a long time: What's going to happen with the U.S. dollar as a reserve currency? And I'll conclude there.

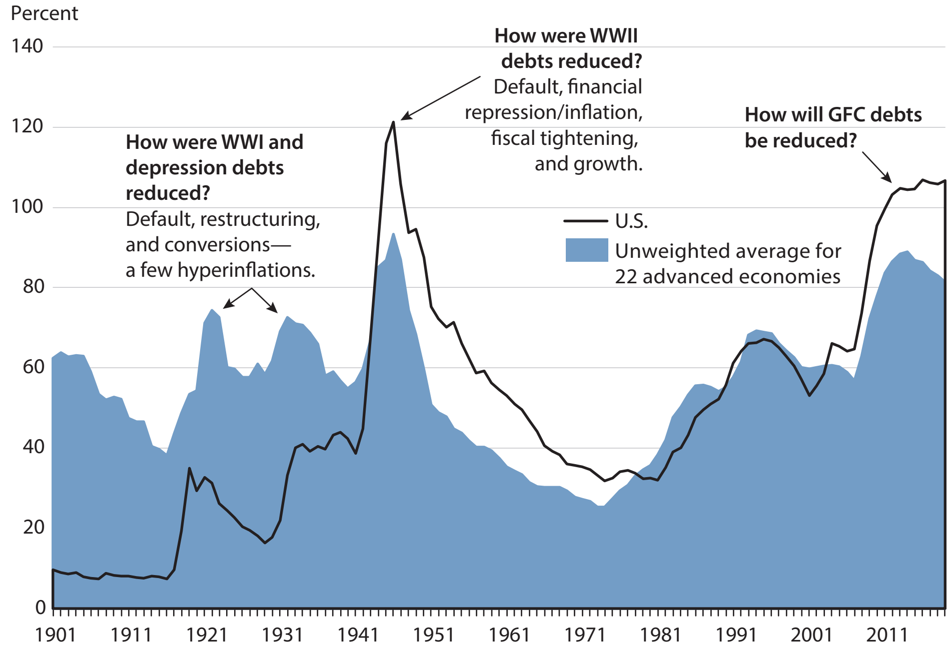
So, on to global risks in the advanced economies. I am going to start where much of the discussion has been recently, which is, of course, on issues relating to trade and globalization but also on issues relating to how much ammunition the advanced economies have in the event of a downturn. Let me start there.

Figure 1—not surprisingly, given the kind of work that I've done in the past on debt—basically shows the level of public debt from 1900 through the present for over 20 advanced economies. I want you to have three takeaways from this graph. Number one is pretty self-evident. Advanced economies as a whole have the highest levels of debt since World War II

Carmen M. Reinhart is the Minos A. Zombanakis Professor of the International Financial System at Harvard Kennedy School.

Figure 1

Advanced Economies: Government Debt as a Percent of GDP, 1901-2019



NOTE: GFC, global financial crisis.

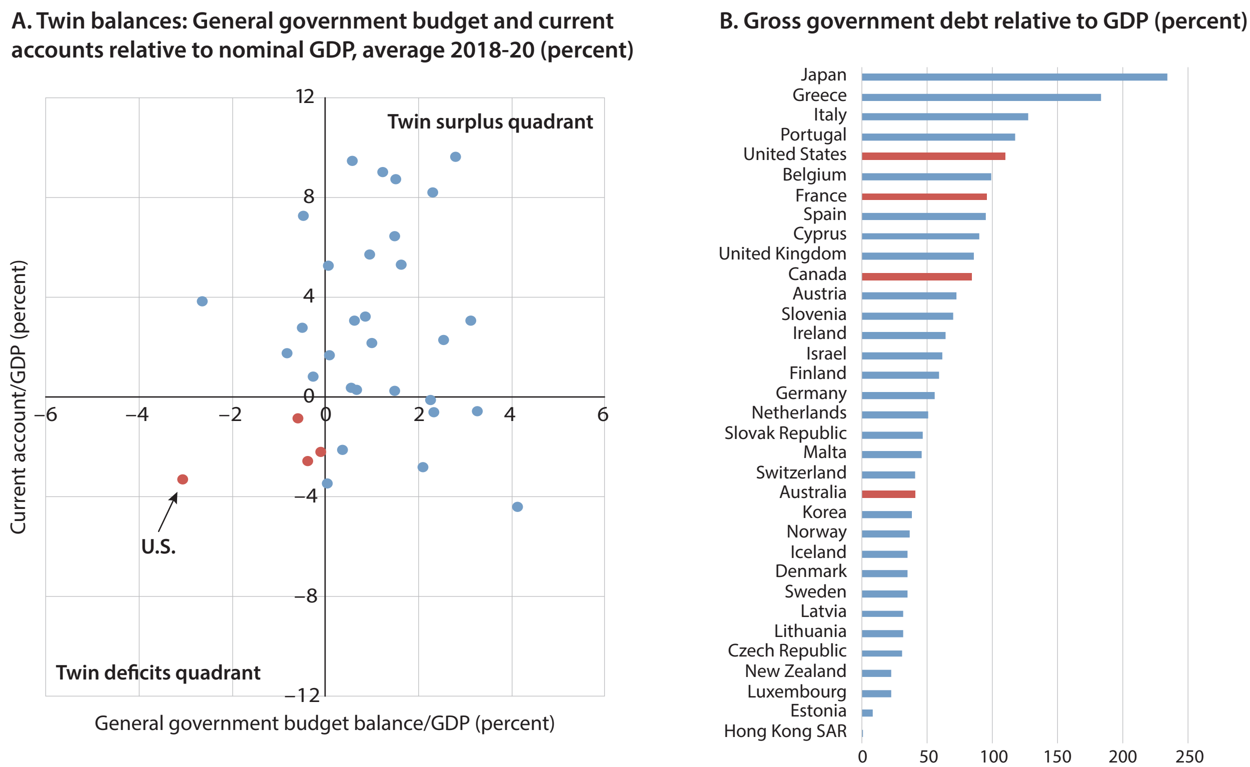
collectively (blue shaded area). Number two, the United States (black line), which had more fiscal space than a number of advanced economies on the eve of the global financial crisis, has since had a bigger surge in debt and now has more sustained debt than what we've seen in other advanced economies.

Number three is tricky because you have to use your imagination: It's not what you see that should worry you; it's what you don't see that should worry you. This is strictly on-budget public debt, and therefore any off-balance-sheet items are not included. And two points on off-balance sheet items. At the end of World War II, public debt was the whole story. Private debt had been unwound through the Great Depression and the war, so the private sector was lean and mean. That applied to households. That applied to corporations. The whole story was what you see in this graph. In addition, advanced economies had much younger populations and very limited pension liabilities at that time, which you also don't see here.

The point that I am making is that, if Figure 1 already highlights that public indebtedness is a limiting factor (a limiting factor in how much fiscal space advanced economies have to cope with a downturn), I would add that limiting that fiscal space are further considerations in that private debt. In the context of the United States, I think there are some concerns on the corporate side. I will talk about that later. But generally, private debt levels are quite high in the advanced economies, certainly very high by postwar measures. And pension liabilities are understandably an issue unlike ever before, because of the aging structure of the popula-

Figure 2

Twin Deficits



NOTE: Red dots and red bars indicate countries with twin deficits.

SOURCE: IMF and World Economic Outlook.

tion and the pension liabilities we've accumulated. So bottom line: Fiscal space is a lot more limited, notwithstanding many arguments out there that debts don't matter and deficits don't matter.

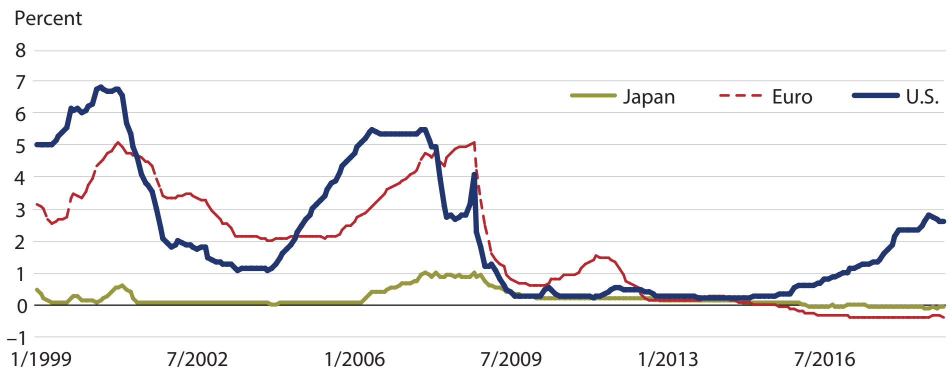
Note again the solid line in Figure 1. As I've mentioned, the United States has less fiscal space than the other advance economies because of the growing U.S. debt. The Congressional Business Office has done studies that basically show that in about a decade's time, even with interest rates remaining where they are or even with interest rates moving lower by about 50 basis points, the United States still has a debt sustainability issue arising—so bear that in mind.

I would also like to point out for the United States that, if we were an outlier in Figure 1 in our accumulation of debt, we're also an outlier in how quickly we are adding both public debt and external debt to our balance sheets. The bottom-left quadrant of Panel A of Figure 2 shows countries with twin deficits. Basically, a twin deficit means you have a current account deficit—you're borrowing from the rest of the world—and a fiscal deficit.

Former Federal Reserve Chairman Bernanke for many, many years talked about the saving glut. The saving glut is basically what you see here on the left side—the deficits. The saving

Figure 3

Three-Month LIBOR Interest Rates



SOURCE: FRED®, Federal Reserve Bank of St. Louis.

Table 1

Monetary Policy "Space"?

Recession	Starting federal funds rate (percent)	Lowest federal funds rate (percent)	Cumulative cut (percentage points)
1990	8.25	3.0	5.25
2001	6.50	1.0	5.50
2007	5.25	0	~5.25

SOURCE: Federal Reserve.

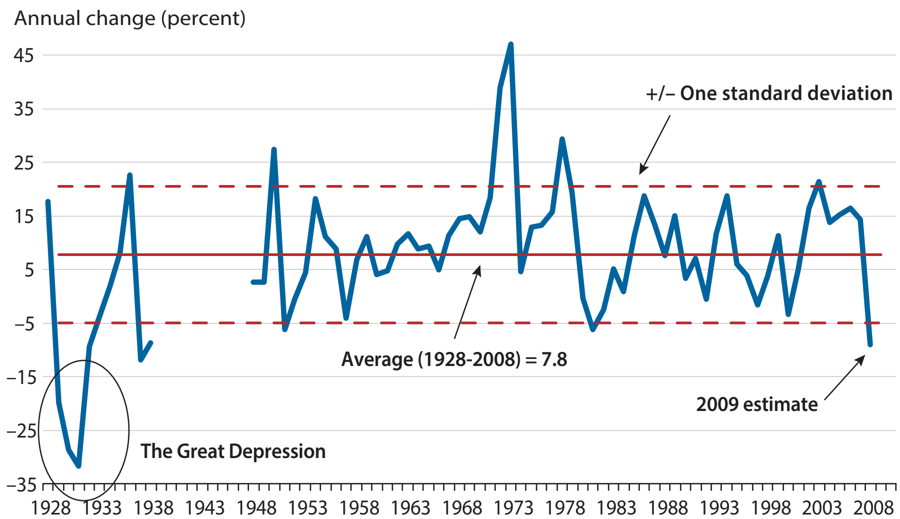
glut basically amounts to China saving a lot. We don't save as much. Surpluses in Asia and Germany also are offset by our deficit. That's an old story. Certainly, it's a story that carried weight in the 1980s. It began in the 1990s, it continued in the 2000s, and it has continued to the present.

What is relatively new is that in addition to the old flow problem, we now have more of a stock problem, meaning we are adding debt when our relative standing in terms of global indebtedness has notched up considerably. I will return to this issue later when I talk about what we can expect over the medium term for the U.S. dollar as a reserve currency.

Again, fiscal space is much more limited now for the advanced economies than it was at the time of the global financial crisis. I would argue that monetary policy space is also much more limited for obvious reasons. As shown in Figure 3, Japan has had negative interest rates for some time. Europe has had negative interest rates for some time. And for the United States, which is the outlier with positive interest rates, in the past, the average decline in the federal funds rate to combat recession had been 600 basis points (Table 1). This is something that we

Figure 4

Global Annual Export Growth, 1928-2009



SOURCE: Reinhart and Rogoff (2009).

are not capable of delivering at the moment. So I think a real risk—and I will later conclude on this point—is that the advanced economies, collectively, are seriously constrained in terms of policy tools to deal with bad shocks. And that is any bad shock.

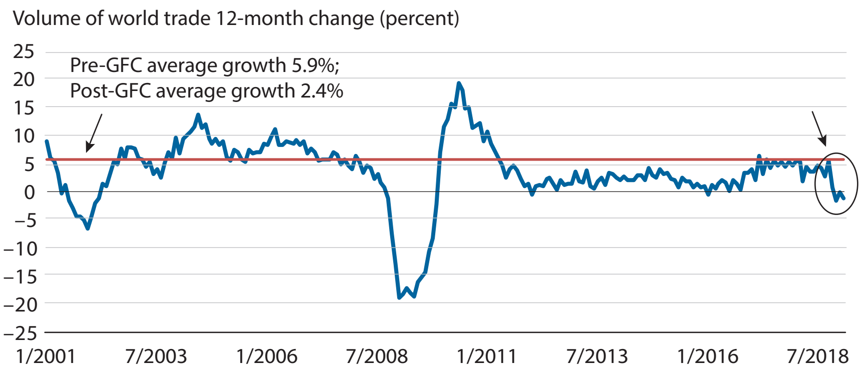
Let's turn to a shock that has been very much in the press—the trade wars. Figure 4 shows global annual export growth from 1928 to 2009. The figure ends in 2009 because it is taken from my book with Ken Rogoff. What are people worried about? Well, people are worried about a replay. In one potentially bad scenario, they're worried about a replay of the aftermath of the Smoot-Hawley tariffs and the trade war of the 1930s, which produced that record contraction in global trade.

Are we there? Are we close to there? What's going on with globalization? Well, let me make a couple of points about globalization. Actually, I wrote about this in *Project Syndicate* years ago. The peak in globalization was in the year before the crisis. Since the crisis, we've been moving toward a lower level of global growth in terms of trade (Figure 5). If you look at the decade before the global financial crisis, average trade growth volume was about 6 percent. In the decade after the crisis, it was less than half of that. This is not unique to the post-global financial crisis experience. This is not the first era of globalization we've had.

I think people don't realize that in the late 1800s to early 1900s, before World War I, we had a very globally integrated capital and goods and service system, albeit limited by the technology at the time. But that globalization was shot to pieces, first by World War I, then by the Depression, and certainly by World War II. Although the financial crisis did not have the extent of drama that the two world wars and major depression produced, it did put a big dent in global trade: It made countries running a current account deficit—such as Spain, Greece,

Figure 5

Global Trade Now in Negative Territory: Will This Time Be Different?



NOTE: GFC, global financial crises.

SOURCE: CPB Word Trade Monitor, February 2019.

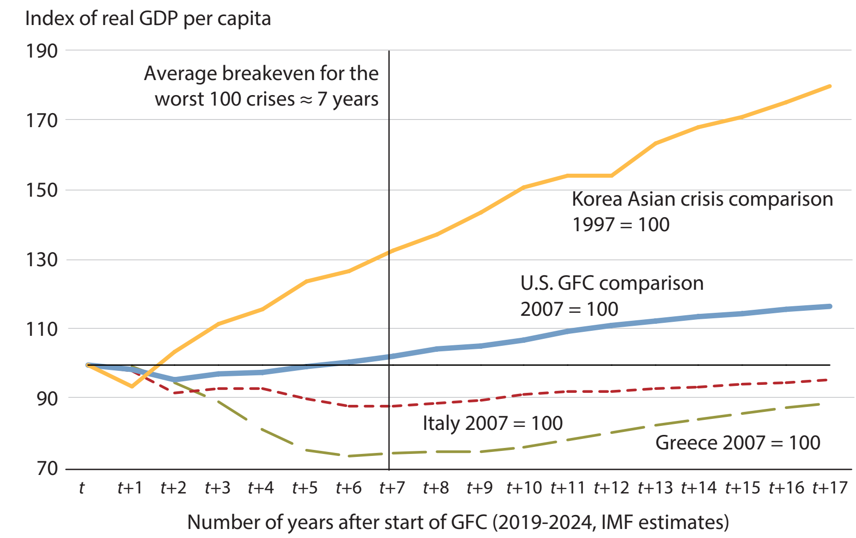
Italy, Ireland, and others—realize that you can't finance a current account deficit from the rest of the world. So you have to watch. You have to import less. You have to look more to home. I think the issue of the rise of home bias dates back to the global financial crisis. I would note that Brexit was another major blow, and more recently, of course, what we're seeing in the trade wars is yet another.

What, in a nutshell, do I take away from the trade wars? Well, I found that my first assessment was completely wrong. If you had asked me in 2018 what I thought, I would have said the trade wars would be resolved a lot quicker, that it would be more of a NAFTA-type situation with a swifter resolution. It wasn't. It isn't. And it's my view now that it's not likely to be because what I've seen is that, over the course of this period, it's become not just about trade, but about geopolitical issues, about security issues and all kinds of issues that are unlikely to be resolved entirely with a handshake. So I think because of the electoral cycle in the United States, we are going to get some news, some deliverables on trade, but not a resolution. And certainly, I'm not looking for a return to the pre-crisis globalization era.

Let's continue on our global tour of advanced economy risks. I just returned from Europe. I was in Paris giving a talk there last week, and always the question is, what do you think are the weak points? Well, I think one can't talk about the next crisis in Europe without really making the point that the previous crisis hasn't been resolved for all of Europe. If you look at Figure 6, two things stand out: Financial crises produce a recession, a drop in per capita GDP, but ultimately countries recover. The recovery is pretty dramatic for Korea, but for Italy and Greece, it's nonexistent. If you take the level of Greek GDP or Italian per capita GDP today, it is below what it was in 2007. And if you take the International Monetary Fund (IMF) projections out to 2024, even by 2024, per capita income in Greece and in Italy will still be below what it was in 2007. That is a pocket of weakness that I think will continue to be a source of both tension and recurring bouts of global uncertainty when we talk about Europe.

Figure 6

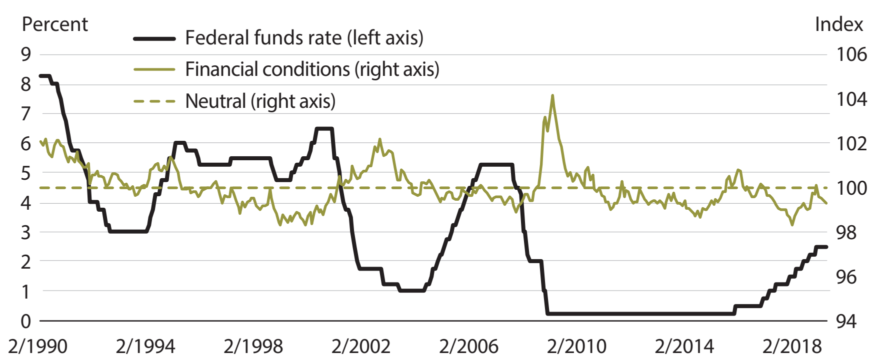
Post-Crisis Recovery (or Lack Thereof in Southern Europe): Unresolved Debt Overhangs



NOTE: GFC, global financial crisis.

Figure 7

Federal Funds Rate and Financial Conditions



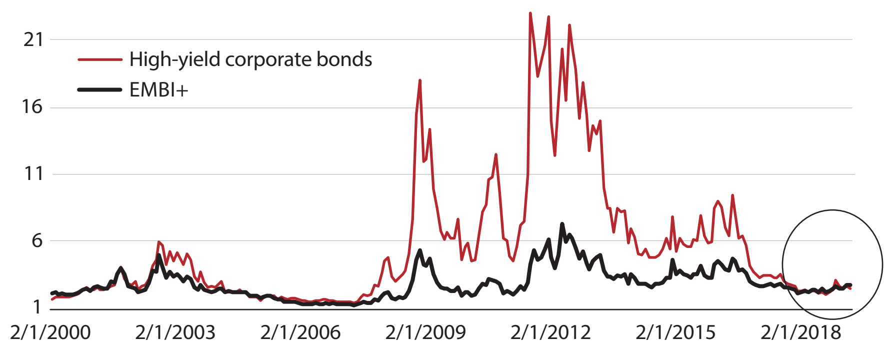
SOURCE: Federal Reserve and Goldman Sachs, accessed via Bloomberg, May 13, 2019.

I say "global uncertainty" because if you go back to 2018, every time the new Italian government made an announcement about possibly leaving the euro or contemplating not servicing debt, those comments translated into an appreciation in the dollar. A depreciation in the euro and an appreciation in the dollar is bad news for emerging markets that have dollar debts. It is a global shock. Let's now move on to yet another but very different type of risk.

U.S. financial conditions remained fairly accommodative through 2018 (Figure 7, black line). Look, for instance, at the financial conditions index that Goldman Sachs publishes

Figure 8

Yield Ratios: Risky Yields Relative to Corporate Debt Yields



NOTE: EMBI+, J.P. Morgan Emerging Bonds Index.
SOURCE: Bloomberg and author's calculations.

(Figure 7, green line). I do not suggest that this is, by any means, the be-all end-all liquidity measurement of financial conditions. But by and large, the point I'm making here is that financial conditions, on the whole, up until the end of 2018 were relatively accommodated. And that, in turn, also contributed—in a world of low yields—to the eternal search for yield. I have done work going back to 1815 on the search for yield (Reinhart, Reinhart, and Trebesch, 2016): The search for yield is eternal. And it drove investors into high-yield corporate debt and, notably, comparatively newer instruments like collateralized loan obligations (CLOs).

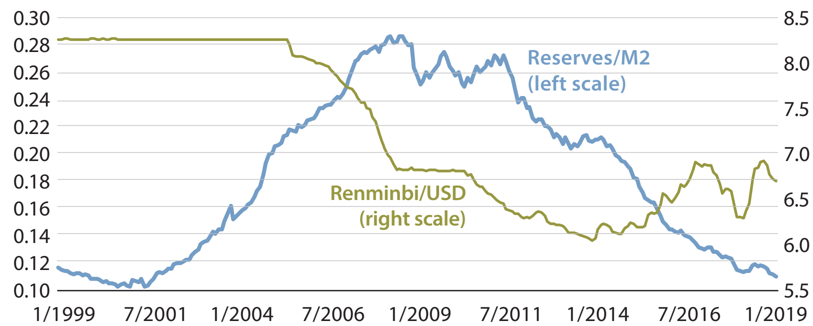
I think we should pause a minute and think about what kinds of risk the honeymoon that we're seeing or have seen in these episodes is bringing to the table. Concretely, CLOs have some similarities to the mortgage-backed security problem in that they're not only attracting local interest but also global interest. In other words, foreign banks: Japanese banks and European banks are coming into the CLO market, which also potentially means that if that market sours, there will be global spillovers into other markets. Let's be clear: I'm not suggesting comparable magnitudes, but magnitudes in some dimension reminiscent of 2008-09.

I always remind people that German banks did not get into trouble in 2009 because they had a real estate bubble in Germany. They did not have a real estate bubble. They got into trouble because they had bought U.S mortgage-backed paper. So there's some scope there for an international contagion dimension.

There is also a worrisome trend: The quality of the borrowers and the quality of the covenants of these loans have been deteriorating. And that is also reminiscent of the run-up to the global financial crisis in which the earlier tranches of the mortgage pools were better quality than the later tranches. Just to be clear, what I'm saying is that, historically, emerging market yields—emerging markets, high-risk debt—have moved together with corporate high-yield debt (Figure 8). What we have seen in the last year and a half is that emerging market yields on the whole went higher while—at the same time—corporate yields went lower.

Figure 9

Foreign Exchange Reserve Ratios and the Currency



NOTE: USD, U.S. dollar.

SOURCE: FRED®, Federal Reserve Bank of St. Louis.

So the question to you is this: What does this mean? Are we overpricing—overestimating—the risks in emerging markets, or are we underestimating the risks in the corporate sector? The arguments I've made thus far is the latter, and that's the convergence that I'm showing in Figure 8.

Let's turn to emerging markets, specifically China, which by almost any metric is the size of the U.S. in the global economy. So we're talking about the world's second-largest economy and, depending how you measure it, in some instances the same size as the U.S. economy. The same points that I made about advanced economies having more-limited ammunition, I am going to make about China—and not about fiscal policy but about monetary policy.

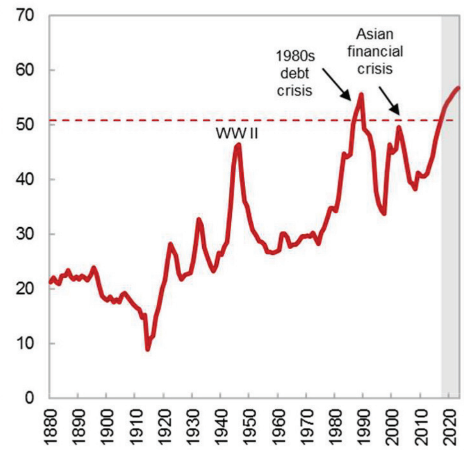
Let me clarify. One of the concerns we've had when discussing the potential for a global slowdown is not just that the United States appears to be slowing or that Europe appears to be slowing, but that China also appears to be slowing—and slowing big-time. We should be concerned, again, because China is the world's second-largest economy. And a level of comfort is usually drawn from the idea that it can provide stimulus. And indeed, China is providing fiscal stimulus to the tune of about 1.75 percent of GDP. However, I don't think it's reasonable to expect China to provide fiscal stimulus and monetary stimulus, which for them is credit creation—meaning providing accessible credit to the corporate sector, to exporters, to banks, and so on.

Let me explain. In 2008-09, China really did record—by almost any metric—fiscal stimulus and monetary stimulus. At that time, however, China was growing double digits. They had large capital inflows. They were accumulating U.S. Treasuries. They were accumulating reserves. They were trying to lean against the wind to avoid a renminbi appreciation. That is not where they are now.

Figure 9 shows international reserves converted into renminbi, divided by M2 (the blue line). What this indicator shows is that China has gone through the full phase of the capital flow cycle. (This indicator goes back to work that Graciela Kaminsky and I [Kaminsky and

Figure 10

General Government Debt: Emerging Market and Middle-Income Economies, 1880-2023



SOURCE: Bredenkamp et al. (2019); https://www.imf.org/en/News/Seminars/Conferences/2018/05/24/sovereign-debt-a-guide-for-economists-and-practitioners.

Reinhart, 1999] did on indicators of the capital flow cycle and indicators of financial crises many, many years ago.) China had a surge in inflows associated with the boom, and it is in the outflow phase.

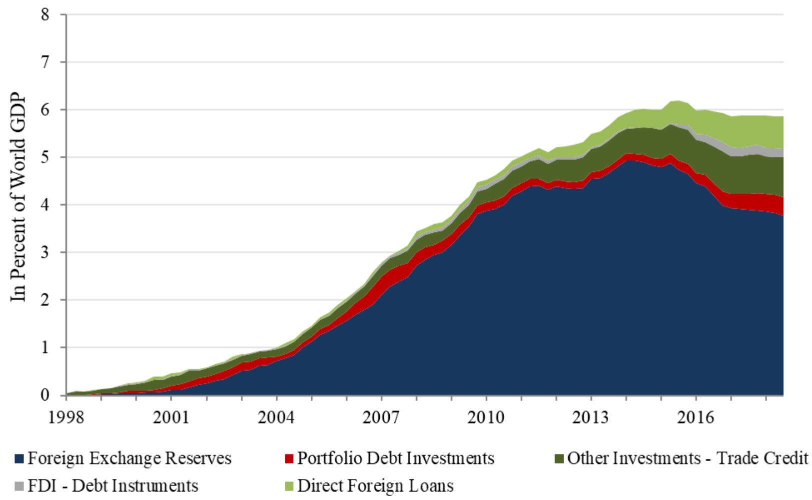
In other words, a country facing capital outflows and trying to maintain a more-or-less stable currency can only do three things. One, they can try to stabilize the exchange rate by losing reserves (i.e., selling their dollar holdings), which China's been doing, intervening to stabilize their renminbi. Two, they can tighten controls, which it's also been doing. This is related to the turmoil in Hong Kong. And three, they can keep tight money, which basically goes to the point that I was making. I don't think China has the ability right now to really engage in very stimulative monetary policy, at least nothing like what it's done in the past.

This is a global risk because China's footprint, as we shall also see, among emerging markets is nothing less than major. We hear a lot about the impact of trade, but China also has a huge impact through finance. And I end my commentary on China by saying that if you do backward exercises, meaning you look at China's trading partners, whether they're commodity producers or other Asian economies that export intermediate goods to China, and you look at how much they've slowed, you would not infer that the remnibi-to-U.S. dollar slowdown (Figure 9, green line) has only been to 6 or 6.5. You would infer that the slowdown is even greater. So I think that that is also a serious—bigger, more protracted—Chinese slowdown that is also a serious headwind to the global economy.

Briefly, I'll also mention two other types of risk now moving out of the big countries and into emerging markets. Figure 10, which is taken from recent work (Bredenkamp et al., 2019) that I did for an IMF conference and volume, shows the indebtedness for emerging markets. Since we are comparing previous episodes, let's choose a relevant previous episode we all are

Figure 11

The Rise of China as a Global Official Creditor, 1998-2018



NOTE: FDI, foreign direct investment.

SOURCE: Horn, Reinhart, and Trebesch (2019).

familiar with: the Asian crisis or the crisis of the 1980s. I would note emerging markets are in more vulnerable territory. Not just Turkey. Not just Argentina. But emerging markets as a whole have slowed down dramatically, largely as a consequence of the slowdown in China and partly, also, as a consequence of dollar strength. These countries tend to have a high share of either corporate or public debt, or both, in U.S. dollars. So a dollar appreciation means higher debt servicing costs, which means more problems. But emerging markets are in more vulnerable territory than they've been in a while.

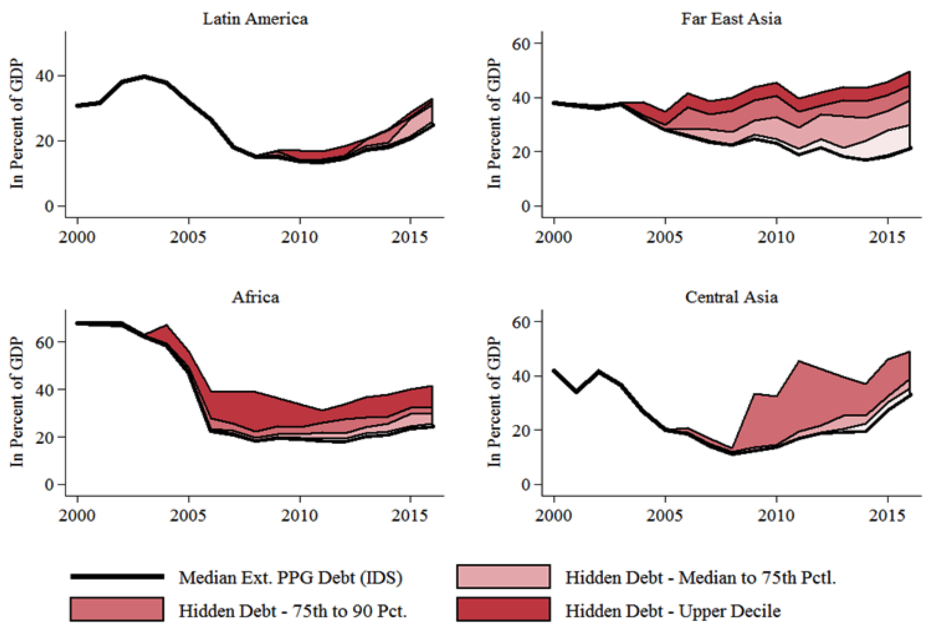
Finally, before talking about the long-term issue, I mentioned that China's role in emerging markets is not just the vast expansion in trade. It's been an expansion in finance. (Figure 11) Right now, Chinese lending to the emerging world is bigger than all such lending from the Paris Club and the IMF combined—the official creditors, all the major advanced economies, that lend bilaterally to emerging markets. China's loans are bigger than all of those and the IMF and the World Bank combined.

Well, what's the problem? To say that the problem is that China's lending is opaque would be an understatement. It is not recorded by the Bank for International Settlement. It is only partially recorded in the World Bank database. It is a thorn in the side of the IMF. The IMF's program with Pakistan—a big to-do—was actually trying to find out how much debt Pakistan owed China. You can't do debt sustainability exercises that are meaningful if you don't know what the outstanding level of debt is.

Hidden debts are a big problem for countries that have borrowed from China. And if you're an investor, you also worry about hidden debts: If you're buying an Ecuadoran or Angolan bond and you're pricing them, thinking that the external debt of that country is, let's say, 40

Figure 12

Total External Debt: Officially Reported (World Bank) and "Hidden Debts" to China, 2000-18



NOTE: Median ext. PPG debt (IDS), median external, public and publicly guaranteed debt to GDP according to the World Bank's International Debt Statistics.

SOURCE: Horn, Reinhart, and Trebesch (2019).

Table 2

Countries Restructuring External Chinese Debt Since 2011

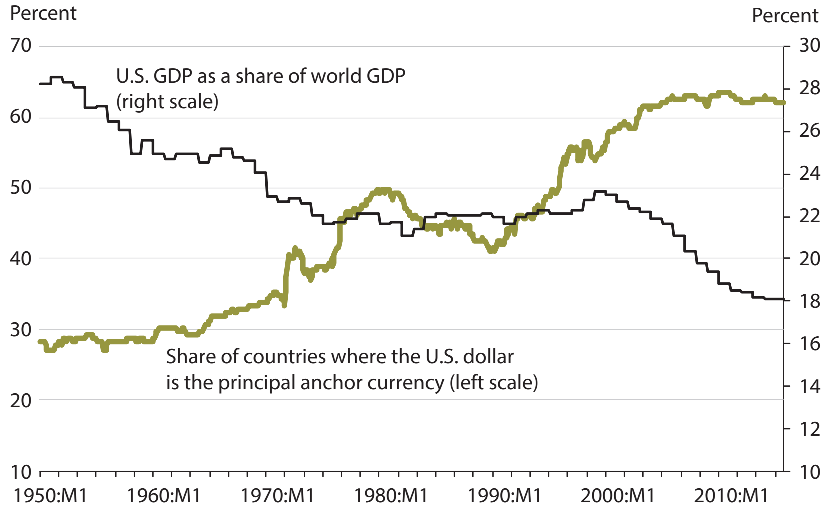
Tanzania	Cuba	Sudan	Ecuador
Bangladesh	Seychelles	Venezuela	Zambia
Cote D'Ivoire	Togo	Ukraine	Sri Lanka

percent, and it's really 60 percent, you have a problem. And if also you don't know who the senior creditors are, you have a problem.

The issue that I'm raising here is based on ongoing work with Christoph Trebesch and Sebastian Horn (Horn, Reinhart, and Trebesch, 2019) (Figure 12). The issue is that these are economies that are not systemic. On the whole, they tend to be small. Collectively, they're not trivial either. And there is a problem of seriously underreported debts. The World Bank database captures only about 50 percent of China's loans to these countries. So with that uplifting note, let me turn to one final point. By the way, it is not a hypothetical that these debts cause problems. If you look at sovereign restructurings, we've had about a dozen sovereign restructurings already of Chinese debt that we know of (Table 2.) There may be more.

Figure 13

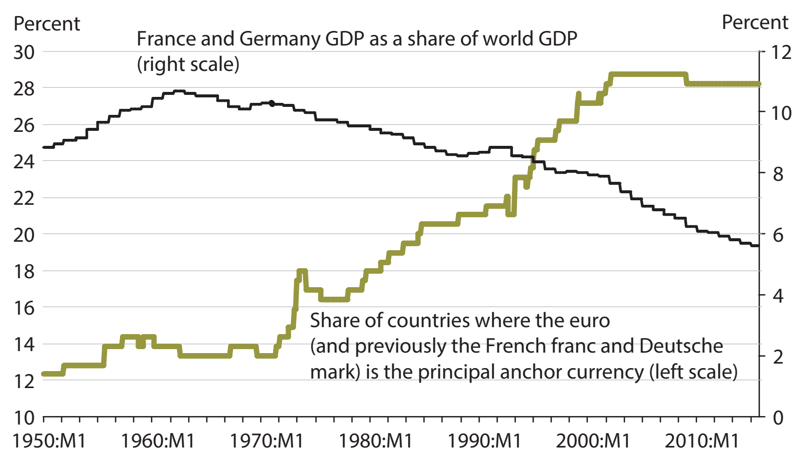
Role of the Dollar and the Global Footprint of the U.S. Economy, 1950-2016



SOURCE: Ilzetzki, Reinhart, and Rogoff (2019).

Figure 14

Role of the French Franc and Deutsche Mark 1950-1998 and Euro 1999-2016



NOTE: Notice the difference in scales in comparison with those for the U.S./U.S. dollar in Figure 13.

SOURCE: Ilzetzki, Reinhart, and Rogoff (2019).

I now turn to my last topic: What about the long horizon? I talked about U.S. debt rising. I talked about the twin deficit problem. I talked about the COB projecting that, even with rates roughly where they are, there is still a debt sustainability problem. Well, let me bring up an old topic—the Triffin dilemma. It arose in the late 1960s when the United States was borrowing heavily to finance the Vietnam War. And at the time, which was still under the Bretton Woods system, countries held dollar reserves and stabilized against the dollar. There was demand for U.S. dollar debt: Remember, countries and central banks do not buy greenbacks. They buy debt. The essence of the Triffin dilemma is that for domestic considerations, you would like to be more circumspect about your debt levels. The external dimension is that if you're the world's reserve currency, you have a lot of rope to hang yourself with: The rest of the world is willing to buy a lot of debt to sustain what may appear like a strictly domestic unsustainable situation.

How did the Triffin dilemma resolve itself last time? It resolved itself with the breakdown of the Bretton Woods system and dollar depreciation. The dollar depreciated versus the Deutsche mark by about 55 percent. Figure 13 is from Ilzetzki, Reinhart, and Rogoff (2019). The solid line shows one of our measures of the demand for dollars, or dollar debts, from the rest of the world; this is the share of countries where the U.S. dollar is the main anchor currency. The dashed line shows U.S. GDP as a share of global GDP. The modern Triffin dilemma, if you will, is that the U.S. share in the global economy is getting smaller while the demand for U.S. assets is getting bigger. (Also see Figure 14.)

How do you reconcile the two? Last time, the reconciliation was, we can say, a devaluation because it was in an era of fixed exchange rates. The question now is, will this mean that once again the equilibrating factor to impose a tax, if you will, on foreign bond holders is a secular depreciation of the dollar? Now, every time you say "secular depreciation of the dollar" at a time of uncertainty, you know you're going to be wrong. You just know you're going to be wrong because every time you say "dollar depreciation" and there's uncertainty, you have to face the flight to quality, the flight to the dollar. Why are we in a situation where the long-term secular trends tell you one thing and in the short run something else happens? I would have to say that at the moment, it's a lack of alternatives.

The fact is that the euro hasn't delivered what everyone hoped it would deliver. There is no liquid euro debt market. You have Italian debt. You have Greek debt. You have a more fragmented system. The renminbi is not a convertible currency. And given the trends that I described earlier for China, it's clear that China has been scaling back on its ambitions to make it an international currency relative to what their ambitions were six years ago.

Is the dollar going to depreciate on a secular basis, or are we going to continue to have dollar appreciation every time global uncertainty pops up? Because if you look at every moment of turmoil, it's usually characterized by a flight into U.S. assets and an appreciating dollar. ■

REFERENCES

Bredenkamp, Hugh; Hausmann, Ricardo; Pienkowski, Alex and Reinhart, Carmen M. "Challenges Ahead," in Ali Abbas, Alex Pienkowski, and Kenneth Rogoff, eds., *Sovereign Debt: A Guide for Economists and Practitioners*. Chapter 9. London: Oxford University Press, 2019; https://doi.org/10.1093/oso/9780198850823.003.0010.

Kaminsky, Graciela and Reinhart, Carmen. "The Twin Crises of Banking and Balance-of-Payments Problems." *American Economic Review*, June 1999, *89*(3), pp. 473-500; https://doi.org/10.1257/aer.89.3.473.

Horn Sebastian; Reinhart, Carmen M. and Trebesch, Christoph. "China's Overseas Lending." Unpublished manuscript, 2019.

Ilzetzki, Ethan; Reinhart, Carmen M. and Rogoff, Kenneth. "Exchange Rate Arrangements in the Twenty-First Century: Which Anchor Will Hold?" *Quarterly Journal of Economics*, May 2019, *134*(2), pp. 599-646; https://doi.org/10.1093/qje/qjy033.

Reinhart, Carmen M.; Reinhart, Vincent and Trebesch, Christoph. "Global Cycles: Capital Flows, Commodities, and Sovereign Defaults, 1815-2015." *American Economic Review*, May 2016, *106*(5), pp. 574-80; https://doi.org/10.1257/aer.p20161014.

Reinhart, Carmen M. and Rogoff, Kenneth S. *The Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, 2009; https://doi.org/10.2307/j.ctvcm4gqx.

Offshoring to a Developing Nation with a Dual Labor Market

Subhayu Bandyopadhyay, Arnab Basu, Nancy Chau, and Devashish Mitra

We present a model of offshoring of tasks to a developing nation characterized by a minimum-wage formal sector and a flexible-wage informal sector. Some offshored tasks are outsourced by the formal sector to the lower-wage informal sector. Productivity improvements in performing offshored tasks in the developing nation increase offshoring, but not necessarily formal-to-informal sector outsourcing, which can cause the developed nation's wage to fall. Productivity improvements in the developing nation's informal sector expand both offshoring and outsourcing, causing the developed nation's wage to rise. When the minimum wage is reduced in the developing nation, the developed nation's wage falls when most of the efficiency gains accrue to the informal sector. (JEL F1)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2020, *102*(3), pp. 237-53.
https://doi.org/10.20955/r.102.237-53

1. INTRODUCTION

This article analyzes developed-to-developing nation offshoring in the presence of a dual labor-market structure in the developing nation. While the developed nation's labor market is assumed to feature flexible wages and full employment, the developing nation is characterized by a dual labor market where a formal and an informal sector coexist. While the formal sector is subject to a minimum-wage regulation, the informal sector is assumed to be able to circumvent that law or the law does not apply to it and pay a lower market-clearing wage. It is also possible that the formal sector circumvents the law by outsourcing to the informal sector or hiring informal or casual workers to perform certain tasks. Consideration of labor-market duality leads to some important departures from the existing literature on trade in tasks, which was pioneered by Grossman and Rossi-Hansberg (2008, GRH hereafter) among others.

As described in Bhagwati and Panagariya (2013), India has over 200 labor regulations that apply to firms in the formal sector. These regulations make labor costs higher than what they otherwise would be and adversely affect the flexibility of firms in responding to shocks.

Subhayu Bandyopadhyay is a research officer and economist at the Federal Reserve Bank of St. Louis. Arnab Basu and Nancy Chau are professors at the Charles H. Dyson School of Applied Economics and Management at Cornell University. Devashish Mitra is a professor at The Maxwell School of Citizenship and Public Affairs at Syracuse University. The authors thank two anonymous referees for their very helpful comments.

In practice, firms find ways of getting around these labor regulations by incurring some costs. For example, Ramaswamy (2003) documents that formal sector manufacturing firms in India are able to circumvent labor regulations by hiring temporary (casual) or contract workers to whom those regulations do not apply. Hasan and Jandoc (2013) show that even in large Indian manufacturing firms with employment over 200 workers, casual or contract workers constitute about 30 percent of total employment. Harris-White and Sinha (2007) provide anecdotal evidence supporting outsourcing of certain activities from formal sector to informal sector firms in India. Sundaram (2015) also provides evidence indicative of outsourcing of relatively labor-intensive activities from formal sector to informal sector firms in India. And, finally, Sundaram, Ahsan, and Mitra (2012, p. 79) provide evidence of "linkages between the formal and informal manufacturing sectors through outsourcing." In addition to the evidence for India, there is also evidence for Mexico showing that about 25 percent of employees of formal firms are informal workers; thus, formal firms are able to avoid many labor regulations (Samaniega de la Parra, 2016).

The paper by GRH is one of the first to model trade in tasks in the context of a developed nation offshoring tasks to a lower-wage nation. The paper's structure is similar to neoclassical competitive models of trade. Accordingly, as in the Heckscher-Ohlin type framework, a reduction in the cost of offshoring has a positive wage effect similar to a productivity increase. This leads to the somewhat counterintuitive result that technological improvements in offshoring that lead to more tasks being offshored can actually lead to a higher wage for labor in the developed nation. This is possible because technological improvements lead to cost savings and scale expansion, and these are reflected in a higher domestic wage at full employment. The paper by GRH focuses on the developed nation, and the nation that performs the offshored tasks is modeled simply as a nation with a fixed wage.

Bandyopadhyay et al. (2020) provide a model for the joint determination of wages in the developed (source) nation that offshores tasks and a developing (recipient) nation that completes the tasks. Within the context of this model, they derive several results that show that while a developed nation may gain from technological improvements in offshoring, the developing nation could lose if the labor-saving effect of technological improvements outweighs the scale-expansion effect. One major issue not considered by Bandyopadhyay et al. (2020) is the importance of the informal sector in developing nations. Indeed, while the formal sector can feasibly be monitored by the government, the informal sector is often out of reach of government regulations. This means that labor standards or minimum-wage laws are hard to enforce in the informal sector, which creates an incentive for firms to outsource some of their tasks to the informal sector. Keeping this duality between the formal and informal sectors in mind, we analyze how technological improvements may impact wages and employment in a simultaneous labor market equilibrium in three markets: the developed nation's labor market, the developing nation's formal sector labor market, and finally the developing nation's informal sector labor market.

We build a model where two nations, which are small in the output market, have a bilateral offshoring relationship in the production of a manufacturing good.[1] As in GRH, competitive firms based in the developed nation produce this good by completing a range of tasks. Some of these tasks are relatively complex and require more labor to be completed in the developing nation, so they are completed in the developed nation, where it is cheaper, while the rest of the tasks are offshored. Among the offshored tasks, intermediate-complexity tasks are completed in the developing nation's minimum-wage formal sector, while the least-complex tasks are completed in its lower-wage informal sector where it is cheaper.[2] This second layer of task allocation is commonly referred to as "domestic outsourcing," which allows formal sector firms to circumvent the minimum wage.[3]

The focus of our general equilibrium model is simultaneous labor market clearing in the developed and the developing nations, where each nation has two sectors, a manufacturing sector and a numeraire agricultural (food) sector.[4] In the developing nation, there is a dual labor market characterized by a rigid-wage formal manufacturing sector and a common flexible wage in the informal manufacturing sector and the agricultural sector. Flexible wages characterize the developed nation's labor market. The residual labor supplies of the manufacturing sectors are absorbed by the respective agricultural sectors of the two nations. We primarily analyze how the flexible wages in the two nations are affected by changes in offshoring technology and outsourcing technology. We also analyze how these factors and parametric changes affect other endogenous variables of interest, such as the levels of offshoring and outsourcing and the share of the informal sector in the developing nation's economy.

The comparative static analysis yields some results that depart from the existing literature. For example, while a rise in offshoring productivity raises offshoring, it may reduce the developed nation's wage. This can happen because the developing nation's informal sector wage may rise through offshoring demand effects and also because of the accompanying shift of marginal tasks from the low-wage informal sector to the higher-wage formal sector. As a result, the degree of informality, given by the ratio of informal-to-formal sector manufacturing employment, may fall. Some other results are counterintuitive at first glance. For example, although increased informal sector productivity in the developing nation will raise formal-to-informal sector (formal-informal) outsourcing, it may reduce both the informal sector's wage and the degree of informality in the nation's manufacturing sector. Similarly, while a minimum-wage cut reduces informality, it may actually increase the informal wage.

Section 2 presents the model and the description of the equilibrium. Section 3 presents comparative static analyses. Section 4 concludes.

2. THE MODEL AND EQUILIBRIUM

2.1 The Basic Structure

Consider two nations, a developed nation F and a developing nation H. There are two homogeneous goods, a numeraire manufactured good and food. We assume that the two nations are small in the output market, so the prices of both goods can be set at unity, without loss of generality. The output levels of the manufactured good and food in nation F are denoted by x^* and y^*, respectively. Nation H also produces food, for which the output level is denoted by y, and workers in nation H may also perform tasks offshored by nation F's manufacturing sector. For simplicity, we assume that all of the manufacturing sector's activity in H is completion of the tasks offshored by F.

Following GRH, we assume that production of a unit of x^* requires a continuum of labor tasks $i \in [0,1]$ to be performed either in H or in F. Labor is the only input used to perform the required tasks. While each task i requires a unit of labor in F, the same task requires $\beta t(i) > 1$ units of labor in H, where β is a general technology parameter and $t(i)$ is the part of technology specific to task i in nation H. Tasks that are more complex and require more labor to complete in the developing nation are indexed by higher values of i. Therefore, by construction $t'(i) > 0$.

Developing nations are often characterized by a dual labor-market environment, where a formal manufacturing sector coexists with (i) an informal manufacturing sector and (ii) the food (agricultural) sector. The formal manufacturing sector features large and well-organized firms bound by laws and regulations: They are required to pay corporate income taxes, get import licenses, have labor unions, etc. The informal manufacturing sector and the agricultural sector are usually characterized by small firms or farmers in rural settings, respectively, where labor laws and regulations do not apply or are not enforceable (because of prohibitive monitoring costs). Accordingly, we first assume that there is a minimum wage in the formal manufacturing sector and a flexible wage in the informal manufacturing sector, where the latter conducts the simplest of manufacturing tasks and is characterized by perfect labor mobility with the agricultural sector. Second, we assume that completion of tasks in the informal sector involves some additional costs. These costs may arise because of a lack of infrastructure that allows the simplest tasks to be transported to the informal sector or the inferior production technology that characterizes the informal sector. Furthermore, to the extent that the informal sector has more infrastructure constraints, such as unreliable electricity, worker productivity in the sector suffers. Since higher values of i represent more-complex tasks, the labor required to outsource from the formal to the informal sector is assumed to be an increasing markup over the labor required to complete the task in the formal sector. This markup is $\tilde{\beta}\tau(i)$, where $\tilde{\beta}$ is a general informal sector technology parameter and $\tau'(i) > 0$ captures that the informal sector is less technologically advanced and thus has increasing difficulty in completing more-complex tasks. The labor required to complete task i in the informal sector is then $\tilde{\beta}\tau(i)\beta t(i)$.

Denoting land and labor by T and L, respectively, the constant-returns-to-scale (CRS) production function for food in nation H is $y = G(L_y,T)$, where L_y is labor used in H's agricultural sector. Similarly, $y^* = G^*(L_y^*,T^*)$ represents nation F's production function for food. Since land is specific to food production and its endowment in each nation is fixed, the CRS production functions for food in the two nations are characterized, respectively, by diminishing returns to labor:

$$(1) \qquad y = G\left(L_y, T\right),\ G_{L_y}\left(L_y, T\right) > 0,\ G_{L_y L_y}\left(L_y, T\right) < 0 \text{ and}$$

$$(2) \qquad y^* = G^*\left(L_y^*, T^*\right),\ G^*_{L_y^*}\left(L_y^*, T^*\right) > 0,\ G^*_{L_y^* L_y^*}\left(L_y^*, T^*\right) < 0.$$

2.2 The Labor Supply in the Manufacturing Sector

Let us denote the developed nation's wage by w^* and the developing nation's wage in the agricultural sector as w. Recalling that output prices are fixed at unity, competitive profit-maximization conditions in the agricultural sectors in nations H and F are $w = G_{L_y}(L_y, T)$ and $w^* = G^*_{L^*_y}(L^*_y, T^*)$, respectively. Inverting these functions and suppressing T and T^* from the functional forms, we obtain the respective labor demand functions in H and F as

(3) $$L_y^d = L_y^d(w),\ L_y^{d\prime}(w) < 0 \text{ and}$$

(4) $$L_y^{*d} = L_y^{*d}\left(w^*\right),\ L_y^{*d\prime}\left(w^*\right) < 0.$$

Given the respective labor endowments $\bar{L}$ and $\bar{L}^*$ of nations H and F, the labor supply functions for the manufacturing sectors of nations H and F are respectively given by[5]

(5) $$L(w) = \bar{L} - L_y^d(w),\ L'(w) > 0 \text{ and}$$

(6) $$L^*\left(w^*\right) = \bar{L}^* - L_y^{*d}\left(w^*\right),\ L^{*\prime}\left(w^*\right) > 0.$$

2.3 Offshoring to the Developing Nation: Formal-Informal Task Allocation

We assume that technology in the agricultural sectors and endowments in the two nations are such that the developed nation's wage w^* exceeds the developing nation's minimum wage $\bar{w}$ in its formal manufacturing sector.[6] Labor mobility between H's informal sector and agricultural sector equalizes the wage between these sectors at w. Although w^* and w are endogenous, labor-allocation decisions are best explained for a given vector of wage rates $(w^*, \bar{w}, w)$. Any task i can be completed by a unit of labor in F at a cost of w^*. This same task can be completed in nation H's formal sector at a lower wage rate $\bar{w}$, albeit with a greater labor requirement $\beta t(i) > 1$. The cost of completing this task in H's formal sector is $\bar{w}\beta t(i)$. As i goes to zero, we have tasks that are less complex and the labor cost of completing these tasks in the developing nation are small enough such that $\bar{w}\beta t(i) < w^*$ and hence the developed nation offshores these tasks. On the other hand, as i goes to 1, we assume that the tasks require sufficiently more labor to be completed in the developing nation such that $\bar{w}\beta t(i) > w^*$, so the tasks are completed in the developed nation. Given continuity and monotonicity of the underlying functions, the marginal offshored task is denoted by I, where

(7) $$\bar{w}\beta t(I) = w^* \Leftrightarrow t(I) = \rho_I \Rightarrow I = I\left(\rho_I\right),\ I'\left(\rho_I\right) = \frac{1}{t'(I)} > 0,$$

where $\rho_I = w^*/(\beta\bar{w})$ is the effective relative factor price of completing a task in the developed nation. Thus, tasks in the range $i \in [0, I]$ are offshored, while the remaining tasks $i \in [I, 1]$ are completed in the developed nation. Next, notice that for the minimum wage to be binding, the informal sector of the developing nation must have a lower wage w. The least-complex offshored tasks (i.e., as i goes to zero) can be performed in the informal sector at a lower cost $w\tilde{\beta}\tau(i)\beta t(i)$ than in the formal sector, where the cost is $\bar{w}\beta t(i)$. This is true for all tasks where

$w\tilde{\beta}\tau(i) < \bar{w}$. On the other hand, the most-complex offshored task (i.e., $i = I$) is such that the high labor requirement dominates the wage advantage of the informal sector, such that $w\tilde{\beta}\tau(I) > \bar{w}$. Thus, task I is completed in the developing nation's formal sector. The marginal task outsourced from the formal to the informal sector is J, where

$$(8) \qquad w\tilde{\beta}\tau(J) = \bar{w} \Leftrightarrow \tau(J) = \rho_J \Rightarrow J = J(\rho_J),\ J'(\rho_J) = \frac{1}{\tau'(J)} > 0,$$

where $\rho_J = \bar{w}/(\tilde{\beta}w)$ is the effective relative factor price of completing an offshored task in the developing nation's formal sector (relative to the informal sector). Given the assumed continuity and monotonicity of the $\tau(i)$ function, (8) implies that out of the offshored tasks, $i \in [0,J]$ are completed in the informal sector and the remainder $i \in [J,I]$ are completed in the formal sector.

2.4 Equilibrium

Given that the manufacturing good is produced through a CRS production technology where each task requires a unit of labor, $(1-I)$ tasks that remain in the developed nation require $x^*(1-I)$ units of labor when output is x^*. Thus, in the presence of offshoring, the labor demand in the manufacturing sector in the developed nation is $x^*(1-I)$. Labor demand in the agricultural sector y^* of the developed nation is $L_y^{*d}(w^*)$. Thus, using equation (6), the developed nation's labor market-clearing condition where the aggregate demand for labor from the two sectors equals the labor endowment $\bar{L}^*$ is

$$(9) \qquad x^*(1-I) + L_y^{*d}(w^*) = \bar{L}^* \Leftrightarrow x^*(1-I) = L^*(w^*).$$

Let us now consider labor required to complete the offshored tasks $i \in [0,I]$. Notice that an offshored task i performed in the developing nation's formal sector requires $\beta t(i)$ labor units. Furthermore, only tasks $i \in [J,I]$ are completed in the formal sector. Since the labor required to complete these tasks vary in the developing nation, the total labor used for completion of these tasks per unit of output is $\beta\int_J^I t(i)di$ in that nation. Therefore, to produce x^* units of output, the labor required in the developing nation's formal sector is $x^*\beta\int_J^I t(i)di$. Similarly, $\beta\tilde{\beta}t(i)\tau(i)$ is the labor requirement to complete a task i in the informal sector, where tasks in the range $i \in [0,J]$ are completed. Thus, production of x^* units of output leads to an informal sector labor demand of $x^*\beta\tilde{\beta}\int_0^J t(i)\tau(i)di$. Labor demand in the developing nation comes from three sources: the formal manufacturing sector, the informal manufacturing sector, and the agricultural sector. The labor demand in the agricultural sector is $L_y^d(w)$. The developing nation's labor market clears when the aggregate labor demand of these three sectors equals the developing nation's labor endowment such that using equation (5) we have

$$(10) \qquad x^*\beta\int_J^I t(i)di + x^*\beta\tilde{\beta}\int_0^J t(i)\tau(i)di + L_y^d(w) = \bar{L} \Leftrightarrow x^*\beta\left[\tilde{\beta}\int_0^J t(i)\tau(i)di + \int_J^I t(i)di\right] = L(w).$$

The flexible-wage rates of the two nations (w,w^*) adjust to clear their respective labor markets simultaneously.

It is convenient to analyze equations (9) and (10) in the form of relative demand and supply between the two nations. If we take the ratio of the left-hand sides of the second equalities in equations (9) and (10), we get the relative demand for labor in the manufacturing sectors of the two nations. Similarly, the ratio of the right-hand sides of the same equations yields the relative supply of labor in the manufacturing sectors of the two nations. The relative demand-supply equality is[7]

$$\beta\left[\tilde{\beta}\mu(J,I)+\gamma(J,I)\right]=\frac{L(w)}{L^*\left(w^*\right)}, \quad (11)$$

where $\mu(J,I)=\left(\int_0^J t(i)\tau(i)di\right)/(1-I)$ and $\gamma(J,I)=\left(\int_J^I t(i)di\right)/(1-I)$. Notice that $\mu(\cdot)$ can be written as $\mu(J,I)=\left[\left(\int_0^J t(i)\tau(i)di\right)x^*\right]/\left[(1-I)x^*\right]$, which is the parameter-adjusted labor demand of the developing nation's informal sector (i.e., informal sector's labor demand divided by $\beta\tilde{\beta}$) relative to labor demand in the developed nation's manufacturing sector. Similarly, $\gamma(\cdot)$ is the parameter-adjusted labor demand in the developing nation's formal sector (i.e., formal sector's labor demand divided by β) relative to the labor demand in the developed nation's manufacturing sector. When offshoring is high, I is larger, and given J, relative demand $\mu(\cdot)$ has to be higher because the unit labor demand in the developed nation's manufacturing sector (i.e., $1-I$) is lower. Similarly, given J, $\gamma(\cdot)$ is increasing in I because of two effects: (i) the aforementioned effect of a reduction in unit labor demand in the developed nation's manufacturing sector and (ii) the expansion of the upper limit of the range $[J,I]$ of tasks performed in the developing nation's formal sector. Similarly, one can explain the effects of changes in J given I. In reality, both (I,J) change in response to changes in relative prices (ρ_I,ρ_J), as described in equations (7) and (8).

The cost of producing a unit of x^* is the sum of the costs of completing all tasks necessary to produce that unit. The cost of completing ($1-I$) tasks in the developed nation is $w^*(1-I)$, while the cost of completing the offshored tasks in the developing nation's formal and informal sectors are $\bar{w}\beta\int_J^I t(i)di$ and $w\beta\tilde{\beta}\int_0^J t(i)\tau(i)di$, respectively. Noting that the price of good x^* is unity, the zero profit condition for the good is

$$w^*(1-I)+\beta\left[w\tilde{\beta}\int_0^J t(i)\tau(i)di+\bar{w}\int_J^I t(i)di\right]=1. \quad (12)$$

Equations (11) and (12) jointly determine the international equilibrium (w,w^*) at a given minimum wage $\bar{w}$ and for given technology parameters β and $\tilde{\beta}$.

3. COMPARATIVE STATICS

The offshoring equilibrium is affected by various parameters underlying the model described in the previous section and also by the minimum wage $\bar{w}$. This section explores how the equilibrium is affected by (i) a change in offshoring technology parameterized by β, (ii) a change in $\tilde{\beta}$ reflecting changes in outsourcing technology related to formal-informal outsourcing within the developing nation, and (iii) a change in the minimum wage in the developing nation's manufacturing sector. In particular, we focus on how changes in these parameters or policy variables affect offshoring (from the developed nation to the developing nation) and outsourcing (from the developing nation's formal to informal sector), the wages in the two nations, and the share of the informal sector in the total manufacturing employment of the developing nation. We first derive some equations that apply to all of the aforementioned parameter and policy changes. After that, we analyze offshoring technology and outsourcing technology changes in Sections 3.1 and 3.2, respectively, and finally the effects of changes in the minimum wage in Section 3.3. Propositions 1, 2, and 3 correspond to Sections 3.1, 3.2, and 3.3 and summarize the findings of each of these subsections.

Let us define the share of the informal sector employment in total manufacturing sector employment in the developing nation as $\delta = \dfrac{\left(\tilde{\beta}\int_0^J t(i)\tau(i)di\right)x^*}{\left(\tilde{\beta}\int_0^J t(i)\tau(i)di + \int_J^I t(i)di\right)x^*}$. Using the definitions of μ and γ above, this reduces to $\delta = \dfrac{\tilde{\beta}\mu}{\tilde{\beta}\mu + \gamma}$. Next, consider the elasticity of the relative demand for labor (see equation (11)) with respect to change in the relative factor price ρ_I, given $(\beta, \tilde{\beta}, \rho_J)$. This elasticity is $\xi^I = \dfrac{d\ln\beta\left(\tilde{\beta}\mu + \gamma\right)}{d\ln\rho_I} = \delta\dfrac{\partial\ln\mu}{\partial\ln\rho_I} + (1-\delta)\dfrac{\partial\ln\gamma}{\partial\ln\rho_I}$ and strictly positive for the following reasons: Given ρ_J, J is fixed and the rise in ρ_I raises I. On inspection of the respective expressions for $\mu(\cdot)$ and $\gamma(\cdot)$ provided below equation (11), it is clear that both these functions are strictly increasing in I. Thus, a rise in ρ_I must raise $\mu(\cdot)$ and $\gamma(\cdot)$, which means that $\xi^I = \delta\dfrac{\partial\ln\mu}{\partial\ln\rho_I} + (1-\delta)\dfrac{\partial\ln\gamma}{\partial\ln\rho_I} > 0$. This elasticity is critical to understanding how the margin for offshoring shifts in response to parametric changes. For example, when β falls, the first-round effect is an increase in the relative price $\rho_I = w^*/(\beta\bar{w})$, which reflects the fact that at a lower β, an offshored task can be performed at a lower wage cost in the developing nation's formal sector. The marginal offshored task increases until the difficulty of transporting the new marginal task offsets the cost savings from technological improvements. As more tasks are offshored, the relative labor demand of the developing nation's formal sector rises, with this effect measured by the term $\dfrac{\partial\ln\gamma}{\partial\ln\rho_I}$. Similarly, the relative labor demand of the developing nation's informal sector rises as the unit labor demand of the developed nation's manufacturing sector falls, with this effect measured by the term $\frac{\partial \ln \mu}{\partial \ln \rho_I}$.

Aggregating these effects by weighting them by the shares δ and $1-\delta$ of the developing nation's informal and formal sectors, respectively, we get the effect on relative labor demand from a rise in the relative factor price ρ_I. This effect is captured by ξ^I above. Similarly, given ρ_I, I is fixed by equation (7) and we can explore the effect of a change in ρ_J on relative labor demand through changes in the outsourcing range $[0,J]$. The effects of changes in ρ_J on relative demand can be expressed by another elasticity represented as a weighted average:

$$\xi^J = \frac{d \ln \beta\left(\tilde{\beta}\mu + \gamma\right)}{d \ln \rho_J} = \delta \frac{\partial \ln \mu}{\partial \ln \rho_J} + (1-\delta)\frac{\partial \ln \gamma}{\partial \ln \rho_J}.$$

This elasticity must also be positive because given ρ_I, I is fixed and the rise in ρ_J raises J, which must in turn raise $\tilde{\beta}\mu + \gamma$.[8] Differentiating (11), we get

$$(13) \qquad \left(\eta^* + \xi^I\right)\hat{w}^* - \left(\eta + \xi^J\right)\hat{w} = \left(\xi^I - 1\right)\hat{\beta} + \left(\xi^J - \delta\right)\hat{\tilde{\beta}} + \left(\xi^I - \xi^J\right)\hat{\bar{w}}.$$

Equation (13) yields an upward-sloping locus in (w, w^*) space because a higher w^* increases offshoring, raising labor demand in nation H so that labor markets of the two nations clear after a suitable increase in the wage rate w. Differentiating (12) we get

$$(14) \qquad \theta^* \hat{w}^* + \left(1-\theta-\theta^*\right)\hat{w} = -\left(1-\theta^*\right)\hat{\beta} - \left(1-\theta-\theta^*\right)\hat{\tilde{\beta}} - \theta\hat{\bar{w}},$$

where θ^* is F's cost share in the production of x^*, θ is H's corresponding cost share of *its* formal sector, and the remainder $(1-\theta-\theta^*)$ is H's cost share of its informal sector. This relationship yields a familiar negative relationship corresponding to the zero-profit condition in the factor price space (w, w^*). Given the output price, a higher wage for labor in nation F can be consistent with zero profit only if the wage for nation H's labor is lower.

3.1 Technological Improvements in Offshoring (Fall in β)

Using equations (13) and (14,) we consider the effects of a change in β (i.e., inverse of labor productivity of offshoring) on w and w^* for a given vector $(\tilde{\beta}, \bar{w})$:

$$\frac{\hat{w}}{\hat{\beta}} = \frac{-\theta^*\left(\xi^I - 1\right) - \left(1-\theta^*\right)\left(\eta^* + \xi^I\right)}{\theta^*\left(\eta + \xi^J\right) + \left(1-\theta^*-\theta\right)\left(\eta^* + \xi^I\right)} \quad \text{and}$$

$$(15) \qquad \frac{\hat{w}^*}{\hat{\beta}} = \frac{\left(1-\theta^*-\theta\right)\left(\xi^I - 1\right) - \left(1-\theta^*\right)\left(\eta + \xi^J\right)}{\theta^*\left(\eta + \xi^J\right) + \left(1-\theta^*-\theta\right)\left(\eta^* + \xi^I\right)}.$$

Proposition 1.[9] *A reduction in* β *leads to*

(i) *an increase in the range of offshoring* $[0,I]$,

(ii) *an increase in w and a decrease in the range of formal-informal outsourcing* $[0,J]$ *if and only if* $\xi^I > \theta^* - \eta^*(1-\theta^*)$,

(iii) an increase in w^* *if and only if* $\xi^I < 1 + \dfrac{(1-\theta^*)(\eta+\xi^J)}{(1-\theta^*-\theta)}$, *and*
(iv) a decrease in δ *if* $\xi^I \geq \theta^* - \eta^*(1-\theta^*)$.

Comment. A reduction in β reflects improved offshoring technology that spurs offshoring of tasks and hence an expansion of the range $[0,I]$. If an increase in the demand for offshoring raises the informal wage, then there is a greater incentive to complete some tasks in the developing nation's formal sector and the range of outsourcing $[0,J]$ decreases. Finally, if technological improvements do not spur a lot of offshoring, then cost savings that raise scale must raise the demand for labor in the developed nation's manufacturing sector, pushing up w^*. This is similar to the productivity effect on the developed nation's wage noted in both GRH and Bandyopadhyay et al. (2020). The effect of β on the equilibrium share δ is more complicated and is discussed below.

The specific results in Proposition 1 are better understood by digging deeper. The decrease in β has the following effects. First, it raises the relative factor price $\rho_I\left(=\dfrac{w^*}{\bar{w}\beta}\right)$ of completing a task in the developed nation, leading to more tasks being offshored to the formal sector of the developing nation. Second, notice that β is an offshoring cost parameter that reflects the labor required to move a task out of the developed nation and applies to both the formal and informal sectors. Therefore, a fall in β at a given I tends to reduce labor demand in both of these sectors. Finally, lower labor costs tend to drive down unit costs and a competitive equilibrium is restored through the expansion of scale. This last effect tends to raise the developing nation's labor demand. The sum of these three effects determines whether the developing nation's labor demand as a whole rises or falls in response to a fall in β. When $\xi^I > \theta^* - \eta^*(1-\theta^*)$, the demand for offshore labor responds strongly to a change in the effective factor price ρ_I. In this case, the expansionary effect on labor demand dominates and the developing nation's labor market clears when more labor flows into the nation's manufacturing sector from its agricultural sector (y) through a rise in the wage w (recall that the labor supply of the manufacturing sector $L(w)$ is positively sloped). When w rises, the effective relative factor price for the formal sector $\rho_J = \bar{w}/(\tilde{\beta}w)$ must fall. This reduces formal-informal outsourcing J. Turning to the effect on the developed nation's wage, notice that if ξ^I is relatively small, the effect of a shift in labor demand is relatively small (toward the developing nation) and dominated by the scale expansion effect and hence w^* rises. Finally, the relative size of the informal sector δ must fall when $\xi^I \geq \theta^* - \eta^*(1-\theta^*)$ because the relative size is independent of scale, and hence all that matters is the range of tasks that are outsourced from the formal to the informal sector. Since J falls when $\xi^I > \theta^* - \eta^*(1-\theta^*)$, a smaller range of tasks are completed in the informal sector. Even when $\xi^I = \theta^* - \eta^*(1-\theta^*)$, the relative size of the informal sector must fall because J remains unchanged but I rises, which means that a higher fraction of offshored tasks are now completed in the formal sector. ■

3.2 Technological Improvements in the Informal Sector (Fall In $\tilde{\beta}$):

The effect of a rise in informal sector productivity (i.e., fall in $\tilde{\beta}$) can be obtained by using equations (12) and (14):

$$\frac{\hat{w}}{\hat{\tilde{\beta}}}=\frac{-\theta^{*}\left(\xi^{J}-\delta\right)-\left(1-\theta^{*}-\theta\right)\left(\eta^{*}+\xi^{I}\right)}{\theta^{*}\left(\eta+\xi^{I}\right)+\left(1-\theta^{*}-\theta\right)\left(\eta^{*}+\xi^{I}\right)} \text{ and}$$

$$\frac{\hat{w}^{*}}{\hat{\tilde{\beta}}}=\frac{-\left(1-\theta^{*}-\theta\right)(\delta+\eta)}{\theta^{*}\left(\eta+\xi^{I}\right)+\left(1-\theta^{*}-\theta\right)\left(\eta^{*}+\xi^{I}\right)}. \tag{16}$$

Proposition 2. *A reduction in $\tilde{\beta}$ leads to*

(i) an increase in the range of offshoring $[0,I]$ *and also an increase in the range of formal-informal outsourcing* $[0,J]$,

(ii) an increase in w if and only if $\xi^{J}>\delta-\dfrac{\left(1-\theta^{*}-\theta\right)\left(\eta^{*}+\xi^{I}\right)}{\theta^{*}}$,

(iii) an increase in w^{*}, and

(iv) a decrease in δ if the $\tau(i)$ schedule is relatively steep at $i=J$.

Comment. A reduction in $\tilde{\beta}$ reflects an improved technology of outsourcing from the formal to the informal sector. The first-round effect should be an increase in the range $[0,J]$ of tasks performed in the informal sector. However, the cost reduction for firms spurs scale expansion and raises demand for labor in the developed nation, raising w^{*} and spurring more offshoring. Thus, the range of offshoring $[0,I]$ rises. If $\tau(i)$ is steep, then there is not much scope for increasing outsourcing, hence share δ falls. The effect on w is more nuanced and explained in more detail in the discussion that follows.

At the initial w, the fall in $\tilde{\beta}$ raises the effective factor price $\rho_J=\bar{w}/(\tilde{\beta}w)$ of completing the tasks in the formal sector compared with the informal sector, which shifts more tasks to the informal sector (i.e., J rises). However, because of the decline in $\tilde{\beta}$, each informal sector task requires less labor, which creates a cost reduction at the initial equilibrium that leads to reallocations that increase scale. The scale expansion drives up labor demand in the developed nation, raising w^{*} and hence $\rho_I=w^{*}/(\beta\bar{w})$. Thus, more tasks are offshored. There are different opposing effects on demand in the informal sector. First, labor demand in the informal sector decreases due to greater efficiency from a lower $\tilde{\beta}$. On the other hand, increased offshoring, increased outsourcing of tasks to the informal sector, and scale expansion all suggest an increase in the labor demand of the manufacturing sector of the developing nation. When the offshoring and outsourcing elasticities (ξ^{I},ξ^{J}) are relatively large, the inequality in part (ii) of Proposition 2 is more likely to be satisfied, and the expansionary effects dominate the contractionary effect of labor-saving technological improvements. In this case, aggregate labor demand of the manufacturing sector of the developing nation rises. The labor market clears at a higher wage w, where more labor moves from the agricultural sector to the manufacturing sector in the developing nation. Finally, consider the ratio of formal-to-informal sector labor employment. Suppose $\tau(i)$ is very steep at the initial equilibrium. As $\tilde{\beta}$ falls, there is not much of a change in J because τ rises rapidly to equal the new factor price ρ_J. Without much change in J, there are two effects of a fall in $\tilde{\beta}$, both of which reduce the ratio δ. First, each informal sector task requires less labor, which shrinks this sector's relative employment through the labor-saving effect. Second, as I rises in response to a higher w^{*}, a rigid J means a greater range of tasks $[J,I]$ are completed in the formal sector. This effect also shrinks δ. In other words, unless the $\tau(i)$ schedule is sufficiently flat to allow for an elastic response of J to a rise in ρ_J, the share of informal sector employment is inversely related to informal sector productivity. ■

3.3 The Effects of a Change in the Minimum Wage

If the developing nation's government decides to change the minimum wage, the effects can be analyzed using equations (12) and (14) as follows:

$$\frac{\hat{w}}{\hat{\bar{w}}} = \frac{-\theta^*\left(\xi^I - \xi^J\right) - \theta\left(\eta^* + \xi^I\right)}{\theta^*\left(\eta + \xi^J\right) + \left(1 - \theta^* - \theta\right)\left(\eta^* + \xi^I\right)},$$

$$\frac{\hat{w}^*}{\hat{\bar{w}}} = \frac{\left(1 - \theta^*\right)\left(\xi^I - \xi^J\right) - \theta\left(\eta + \xi^I\right)}{\theta^*\left(\eta + \xi^J\right) + \left(1 - \theta^* - \theta\right)\left(\eta^* + \xi^I\right)}. \tag{17}$$

Proposition 3. *A reduction in* $\bar{w}$ *leads to*

(i) an increase in the range of offshoring $[0,I]$ *and a decrease in the range of formal-informal outsourcing* $[0,J]$,

(ii) an increase in w *if and only if* $\xi^I > \dfrac{\theta^*\xi^J - \theta\eta^*}{\theta^* + \theta}$,

(iii) an increase in w^* *if and only if* $\xi^I < \dfrac{\left(1-\theta^*\right)\xi^J + \eta\theta}{1 - \theta^* - \theta}$, *and*

(iv) a decrease in δ.

Comment. A minimum-wage cut reduces the effective wage of completing the marginal offshored task in the formal sector of the developing nation, which must raise the range of offshoring $[0,I]$. In addition, it also reduces the effective cost of completing the marginal outsourced task in the formal sector compared with the informal sector. Thus, fewer tasks are done in the informal sector, reducing the outsourcing range $[0,J]$ and also the relative size of the informal sector δ. The wage effects are more nuanced and are better understood in the detailed discussion below.

A cut in $\bar{w}$ raises the relative price of completing tasks in the developed nation (i.e., ρ_I) and reduces the relative price ρ_J of completing tasks in the developing nation's formal sector (*vis-á-vis* the informal sector). This expands the offshoring margin I and shrinks outsourcing margin J and has three effects on the informal wage w. First, as the marginal offshored task I rises, demand shifts from the developed to the developing nation, tightening the latter's labor market and exerting upward pressure on w. Second, at a lower minimum wage, more tasks are completed in the developing nation's formal sector, reducing the demand for labor in the informal sector, which has a negative impact on the informal wage. Finally, lowering the unit cost at the initial equilibrium leads to scale expansion, which raises demand in all of the labor markets, exerting upward pressure on all the flexible factor prices. If ξ^I is large relative to ξ^J, the formal-informal reallocation effect (i.e., the second effect) is small and the expansionary effects dominate. The net increase in the manufacturing sector's demand for labor in the developing nation drives up the informal wage w. Finally, the comparative statics effect on w^* is best understood by focusing on how the factor rewards $(w^*,\bar{w},w)$ in the unit cost function may change *vis-á-vis* each other. When $\bar{w}$ falls, given the technology and output price, zero profit requires that at least one of the factor rewards (w^*,w) rises. When ξ^J is large relative to ξ^I, demand shifts disproportionately from the informal to the formal sector in response to the minimum-wage cut. In this situation, there may be a net reduction in the developing nation's manufacturing sector's labor demand, which requires w to fall to clear the market. When w falls, the only possible outcome consistent with a zero-profit equilibrium is a higher w^*. Put differently, if ξ^J is relatively small, then it is possible that the informal wage w rises (as explained above)—to such an extent that even at a lower $\bar{w}$ zero profit can be reestablished only through a fall in w^*. It is easy to check that $\frac{\theta^*\xi^J-\theta\eta^*}{\theta^*+\theta}<\frac{(1-\theta^*)\xi^J+\eta\theta}{1-\theta^*-\theta}$. Using this fact and part (iii) of Proposition 3, we have that if w^* falls, it must be that $\xi^I>\frac{(1-\theta^*)\xi^J+\eta\theta}{1-\theta^*-\theta}\Rightarrow\xi^I>\frac{\theta^*\xi^J-\theta\eta^*}{\theta^*+\theta}$. In turn, using part (ii) of the proposition and the last inequality in the previous sentence, it must be that a necessary (but not sufficient) condition for w^* to fall is a rise in w. In other words, a rising factor reward in the informal sector is what allows the developed nation's wage to fall in spite of the fall in the developing nation's minimum wage. Finally, notice that a larger I and a smaller J in response to a minimum-wage cut imply that fewer tasks $[0,J]$ are completed in the informal sector and a greater range of tasks $[J,I]$ are completed in the formal sector. Thus, the ratio of informal sector employment δ must decline. ■

4. CONCLUSION

This article argues that given the overwhelming importance of the informal sector in many developing nations, it is important to consider the dual labor-market structure that characterizes these nations. It is important not only because the structure is closer to reality, but also because it leads to important differences in the comparative statics responses. While GRH and Bandyopadhyay et al. (2020) both point to developed-nation wage increases due to the productivity effect, we find that the dual labor-market feature can overturn this effect. We see this in Proposition 1 of this article, which notes the possibility of a reduction in the developed nation's wage, in contrast to Proposition 1 of Bandyopadhyay et al. (2020). The reduction is possible because when offshoring elasticity is large, the increase in labor demand in the developing nation can push up costs on two fronts: a higher informal wage and a greater share of the work performed in the relatively higher-cost formal sector. These reallocation effects and factor price changes allow a fall in the developed nation's wage in spite of technological improvements.

Other important factors that are missed by models that do not consider the informal sector is the possibility of purely domestic factors that can raise the productivity of the informal sector. While the direct effect of such changes is a boost in the informal sector, the indirect effect encompasses the offshoring decision as well as the developed nation's wage. With more work being done more efficiently by the informal sector and with the factor price of the formal sector being held constant by the minimum wage, the developed nation's wage s must rise to reflect this efficiency. Such an international effect of a purely domestic technological change is missed by models that ignore the dual labor-market structure. Finally, without a dual labor-market structure, one cannot fruitfully talk about the impact of changes in minimum-wage laws on the vast majority of urban informal workers. Our work shows that while a cut in the minimum wage will shrink the relative size of the informal sector, informal workers can actually be better off because of a rise in the informal wage through the expansionary effects of a minimum-wage cut.

Our agenda for future work on the topic of dual labor markets includes the analysis of the effects of different types of labor standards (including the minimum wage) after allowing for imperfect monitoring of these standards in the formal sector. It is also relevant to look at competing offshoring destinations and how labor standards or the degree of informality in one nation affects other offshoring recipients and possibly their labor standards. Finally, we have abstracted in this article from considerations arising out of terms-of-trade changes in the output market. Interactions between output-market terms of trade and factor-market terms of trade in the presence of informality is another possible avenue for our future work. ■

APPENDIX

A1. PROOF OF PROPOSITION 1

Given $\bar{w}$, the definition of ρ_I in (7), and also (15), we get

(A1) $$\frac{\hat{\rho}_I}{\hat{\beta}}=\frac{\hat{w}^*}{\hat{\beta}}-1=\frac{-\left(1-\theta^*-\theta\right)\left(1+\eta^*\right)-\left(\eta+\xi^J\right)}{\theta^*\left(\eta+\xi^J\right)+\left(1-\theta^*-\theta\right)\left(\eta^*+\xi^I\right)}<0.$$

Equation (A1) implies that a fall in β must raise ρ_I. Therefore, using (7), we have that I must rise when β falls $\left(\frac{dI}{d\beta}<0\right)$. The first relationship in (15) establishes that $\frac{\hat{w}}{\hat{\beta}}<0 \Leftrightarrow \xi^I>\theta^*-\eta^*\left(1-\theta^*\right)$. Notice from the definition of ρ_J that given $(\tilde{\beta},\bar{w})$, $\hat{\rho}_J=-\hat{w}$. Thus, $\frac{\hat{\rho}_J}{\hat{\beta}}=-\frac{\hat{w}}{\hat{\beta}}>0 \Leftrightarrow \xi^I>\theta^*-\eta^*\left(1-\theta^*\right)$. In turn, from (8) we get $\frac{\hat{J}}{\hat{\beta}}>0 \Leftrightarrow \xi^I>\theta^*-\eta^*\left(1-\theta^*\right)$. Now, the second relationship in (15) shows that $\frac{\hat{w}^*}{\hat{\beta}}<0 \Leftrightarrow \xi^I<1+\frac{\left(1-\theta^*\right)\left(\eta+\xi^J\right)}{\left(1-\theta^*-\theta\right)}$. Finally, notice that $\delta=\frac{\tilde{\beta}\mu}{\tilde{\beta}\mu+\gamma}=\frac{\tilde{\beta}}{\tilde{\beta}+\lambda(J,I)}$, where $\lambda(J,I)\equiv\frac{\gamma}{\mu}=\frac{\int_J^I t(i)di}{\int_0^J t(i)\tau(i)di}$. As shown above, when $\xi^I\geq\theta^*-\eta^*(1-\theta^*)$, w will either rise or be constant when β falls. Thus, the marginal task J will either fall or remain constant. The increase in I without any increase in J means that the numerator of the expression for $\lambda(J,I)$ rises, but the denominator remains constant or falls. Thus, $\lambda(J,I)$ must rise, implying that δ must fall when $\xi^I\geq\theta^*-\eta^*(1-\theta^*)$. ■

A2. PROOF OF PROPOSITION 2

The second relationship in (16) shows that the developed nation's wage w^* must always rise when $\tilde{\beta}$ falls. In turn, this means that $\rho_I = w^*/(\beta\bar{w})$ must rise, which implies that I must rise.

Using the first relationship in (16) we get
$\frac{\hat{w}}{\hat{\tilde{\beta}}}+1=\frac{\theta^*(\eta+\delta)}{\theta^*\left(\eta+\xi^J\right)+\left(1-\theta^*-\theta\right)\left(\eta^*+\xi^I\right)}>0 \Leftrightarrow \frac{d\left(w\tilde{\beta}\right)}{d\tilde{\beta}}>0$. Thus, $\rho_J = \bar{w}/(\tilde{\beta}w)$ must rise when $\tilde{\beta}$ falls, which implies that J must rise. Using the first relationship in (16) we find that $\xi^J > \delta - \frac{\left(1-\theta^*-\theta\right)\left(\eta^*+\xi^I\right)}{\theta^*}$ is a necessary and sufficient condition for the informal wage w to rise when $\tilde{\beta}$ falls. Turning to the relative size of the informal sector, recall that $\delta=\frac{\tilde{\beta}}{\tilde{\beta}+\lambda(J,I)}$, where $\frac{\gamma}{\mu}=\frac{\int_J^I t(i)di}{\int_0^J t(i)\tau(i)di}\equiv\lambda(J,I)$. A fall in $\tilde{\beta}$ for a given λ reduces δ. However, since I and J both rise, the direction of the change in λ is, in general, ambiguous. If the $\tau(i)$ schedule is steep at $i = J$, the comparative static change in J will be small. In this event, the denominator for the expression for λ does not change much, but the numerator rises because of a rise in I. Thus, λ rises (assuming that $t(i)$ is not too steep at $i = I$). Therefore, in this case, a reduction in $\tilde{\beta}$ and an increase in λ both reduce δ. If both schedules $t(i)$ and $\tau(i)$ are steep, the offshoring and outsourcing margins do not change much and λ does not change much. However, the fall in $\tilde{\beta}$ reduces δ. Therefore, as long as $\tau(i)$ is sufficiently steep at $i = J$, the informal share δ must fall with a fall in $\tilde{\beta}$. ■

A3. PROOF OF PROPOSITION 3

Using equations (7), (8), and (17) for a given β and $\tilde{\beta}$, we get $\frac{\hat{\rho}_I}{\hat{\bar{w}}}=\frac{\hat{w}^*}{\hat{\bar{w}}}-1<0$ and $\frac{\hat{\rho}_J}{\hat{\bar{w}}}=1-\frac{\hat{w}}{\hat{\bar{w}}}>0$. These imply that a minimum-wage cut must raise ρ_I and reduce ρ_J. In turn, equations (7) and (8) show that I must rise and J must fall. The two inequalities in (17) yield $\frac{\hat{w}}{\hat{\bar{w}}}>0$ if and only if $\xi^I<\frac{\theta^*\xi^J-\theta\eta^*}{\theta^*+\theta}$ and $\frac{\hat{w}^*}{\hat{\bar{w}}}<0$ if and only if $\xi^I<\frac{\left(1-\theta^*\right)\xi^J+\eta\theta}{1-\theta^*-\theta}$. Finally, recall that $\delta=\frac{\tilde{\beta}}{\tilde{\beta}+\lambda(J,I)}$, where $\lambda(J,I)=\frac{\int_J^I t(i)di}{\int_0^J t(i)\tau(i)di}$. As I rises and J falls, the numerator of the expression for λ rises and the denominator shrinks. Thus, λ rises as the minimum wage falls, meaning that δ must fall. ■

NOTES

1 Bandyopadhyay et al. (2020) also present their main results in the context of two nations that are small in the output market. The appendix of their paper considers how the analysis may be extended to large countries and how that can modify their findings.

2 The more-complex tasks will also be more costly to domestically outsource, since greater skills might be required to perform them and skills cannot be fully transferred from the formal to the informal sector. Also, supervision by formal-firm managers of informal sector firms or of casual workers is more difficult. Informal workers, due to the temporary nature of their jobs, have little incentive to acquire skills on the job. For the same reason, their employers have virtually no incentive to invest in their human capital or productivity. Despite the low productivity of informal workers, formal firms transfer some of the relatively simple tasks to them because of the lower informal sector wage.

3 See, for example, Goldschmidt and Schmieder (2017), where "domestic outsourcing" in Germany is analyzed.

4 The small-nation assumption in the output market considerably simplifies the analysis of factor markets by ensuring that any excess supply or excess demand in the output market is absorbed by the world market at fixed international prices. Utility of each nation is entirely determined by national income because output prices are fixed in the indirect utility function of each nation. Of course, national income is endogenous and determined by factor allocation between the two sectors—manufacturing and agriculture. Excess labor supply from agriculture in each nation is absorbed under labor market clearing (modeled) in the offshoring manufacturing sector. Technological change affects allocation of labor both between the sectors and across the nations, and all of this is considered in our analysis. Dropping the small-nation assumption is possible, and following Bandyopadhyay et al. (2020, pp. 222-23) we may pursue this line of inquiry in our future work. The analysis is done in the context of a representative North-South model of offshoring. It may be possible to extend the analysis to several such nations to characterize the global economy. However, such a model is much more complex and beyond the scope of this article.

5 We assume that the technology in sectors (y,y^*) and the endowments $(\bar{L},\bar{L}^*)$ are such that there is an excess supply of labor in the agricultural sector for the relevant range of wages within which a sensible interior offshoring equilibrium (described in the next section) obtains.

6 In our competitive model, each firm hires similarly and the excess labor supply is absorbed at a flexible wage that prevails in the informal sector and in the agricultural sector. In a model of heterogeneous firms (that will require some sort of imperfect competition, such as monopolistic competition with product differentiation) with firm-specific wage negotiations, although the minimum wage may not be binding for all firms, it could be binding for the marginal firm. Our competitive model captures this behavior in a simple and tractable way without invoking a monopolistically competitive framework and the associated complexity.

7 Note that equations (7) and (8) above show that I is entirely determined by $\rho_I = w^*/(\beta\bar{w})$ and J by $\rho_J = \bar{w}/(\tilde{\beta}w)$. If we explicitly note these relationships in equation (11), then we get a relationship between the factor prices $(w,w^*,\bar{w})$ and the technology parameters $(\beta,\tilde{\beta})$. For analytical convenience, we take a slightly different approach, although this aforementioned relationship is at the heart of equation (13) later in the text. Equation (13) captures the labor-market equilibrium in the two nations in the presence of offshoring, outsourcing, and formal-informal duality. On the other hand, equation (14) is derived from the zero-profit condition of competitive firms and also represents a relationship between these factor prices and technology parameters. Equation (14) ensures that a representative firm's scale has to adjust to ensure price-unit cost equality. Equations (13) and (14) together characterize the comparative static effects of parametric changes.

8 Notice that $\frac{\partial\left(\tilde{\beta}\mu+\gamma\right)}{\partial J}=\frac{t(J)\left[\tilde{\beta}\tau(J)-1\right]}{1-I}>0$, because using equation (8), we have $\tilde{\beta}\tau(J)=\frac{\bar{w}}{w}>1$.

9 Proofs of all propositions are provided in an appendix.

REFERENCES

Bandyopadhyay, S.; Basu, A.K.; Chau, N.H. and Mitra, D. "Consequences of Offshoring to Developing Nations: Labor-Market Outcomes, Welfare, and Corrective Interventions." *Economic Inquiry*, 2020, *58*(1), pp. 209-24; https://doi.org/10.1111/ecin.12833.

Bhagwati, J. and Panagariya, A. *Why Growth Matters? How Economic Growth in India Reduced Poverty and the Lessons for Other Developing Countries*. New York: Public Affairs, 2013.

Goldschmidt, D. and Schmieder, J.F. "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure." *Quarterly Journal of Economics*, 2017, *132*(3), pp. 1165-217.

Grossman, G.M. and Rossi-Hansberg, E. "Trading Tasks: A Simple Theory of Offshoring." *American Economic Review*, 2008, *98*(5), pp. 1978-97; https://doi.org/10.1257/aer.98.5.1978.

Harris-White, B. and Sinha, A. *Trade Liberalization and India's Informal Economy*. Oxford University Press, 2007.

Hasan, R. and Jandoc, K.R.L. "Labor Regulations and Firm-Size Distribution in Indian Manufacturing," in J. Bhagwati and A. Panagariya, eds., *Reforms and Economic Transformation in India*. Oxford University Press, 2013, pp. 15-48.

Ramaswamy, K.V. "Liberalization, Outsourcing and Industrial Labor Markets in India: Some Preliminary Results," in S. Uchikawa, ed., *Labor Market and Institution in India: 1990s and Beyond*. Manohar, 2003.

Samaniega de la Parra, B. "Formal Firms, Informal Workers and Household Labor Supply in Mexico." Unpublished manuscript, Department of Economics, University of Chicago, 2016.

Sundaram, A. "The Impact of Trade Liberalization on Micro Enterprises: Do Banks Matter? Evidence from Indian Manufacturing." *Oxford Bulletin of Economics and Statistics*, 2015, *77*(6), pp. 832-853; https://doi.org/10.1111/obes.12082.

Sundaram, A.; Ahsan, R.N. and Mitra, D. "Complementarity between Formal and Informal Manufacturing in India: The Role of Policies and Institutions," in J. Bhagwati and A. Panagariya eds., *Reforms and Economic Transformation in India*. Oxford University Press, 2012, pp. 49-85.

Asset Pricing Through the Lens of the Hansen-Jagannathan Bound

Christopher Otrok and B. Ravikumar

Stochastic discount factor (SDF) models are the dominant framework for modern asset pricing. The Hansen-Jagannathan bound is a characterization of the admissible set of SDFs, given a vector of asset returns. The admissible set provides (i) a test of the asset-pricing model and (ii) information on how to modify the SDF to be consistent with asset returns, neither of which requires solving the model. In this article we use the Hansen-Jagannathan bound to examine asset-pricing implications and to test specific asset-pricing models using bootstrap experiments. (JEL G1, C15, E44)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2020, *102*(3), pp. 255-69.
https://doi.org/10.20955/r.102.255-69

1 INTRODUCTION

An asset-pricing model is typically defined by its stochastic discount factor (SDF). For instance, Mehra and Prescott (1985) used constant-relative-risk-aversion (CRRA) preferences and the SDF in their model was a function of consumption growth. The validity of an SDF is determined by its ability to match the observed asset returns. An early test of an asset-pricing model with CRRA preferences was the Hansen and Singleton (1982) J-test. For U.S. stock and bond returns data, this test typically rejects the model. The J-test tells us whether or not an asset-pricing model has statistically significant pricing errors. It does not provide information on how to modify the SDF to improve the fit. Hansen and Jagannathan (1991) derive a volatility bound (HJ bound) that is based on necessary conditions that an asset-pricing model must satisfy. The HJ bound characterizes the admissible set of SDFs that is consistent with the observed asset returns.

The HJ bound exploits two conditions: (i) the intertemporal Euler equation that connects the price of an asset to the covariance of the asset's payoff with the SDF and (ii) the implication from linear pricing that the SDF be a linear function of payoffs. The asset-pricing model is said to be consistent with the data if the volatility of the proposed SDF (evaluated at the mean SDF) is greater than the volatility implied by the HJ bound. The HJ bound is a lower bound

Christopher Otrok is the Sam B. Cook Professor of Economics at the University of Missouri–Columbia and a research fellow at the Federal Reserve Bank of St. Louis. B. Ravikumar is an economist, a senior vice president, and the deputy director of research at the Federal Reserve Bank of St. Louis.

and, hence, is a necessary but not sufficient condition that an asset-pricing model must satisfy. In other words, the HJ bound provides a "test" of an asset-pricing model based solely on necessary conditions implied by the model.

The HJ bound approach in a sense works backward: Instead of writing down a model, solving it, and then testing it, the HJ bound asks what a valid SDF should look like in the mean-variance space. The HJ bound approach has several advantages. First, the bound is model-free; that is, it is constructed using only observed asset returns. Second, one does not need to solve the nonlinear asset-pricing model. Specifically, there is no need to find a partial equilibrium or a general equilibrium solution to the model. Third, there is no limit on the number of assets used in the construction of the bound. Fourth, the bound is informative on how to modify the SDF in order to be consistent with the data.

In this article we provide a derivation of the HJ bound and then apply the bound to examine a few popular SDFs. The results provide an illustration of the equity premium puzzle. We then check the robustness of the resolutions of the puzzle with a bootstrap experiment. Our bootstrap results indicate that minor variations in asset return moments and consumption moments can yield large variations in the distance between an SDF's volatility and the HJ bound. We conclude with some implications for business cycle models.

2 THE HANSEN-JAGANNATHAN BOUND

For frictionless asset-pricing models, Hansen and Jagannathan (1991) showed that the volatility of the SDF that satisfies the representative consumer's Euler equation must exceed a lower bound that is a function of only asset returns. The derivation of the HJ bound is presented here purely for completeness. (In the appendix, we derive the Sharpe-ratio version of the HJ Bound; see also Ljungqvist and Sargent, 2018.)

Let R denote the $n\times 1$ (gross) return vector of risky assets. Consider an SDF m that prices the n assets according to

$$E_t\left(R_{t+1}m_{t+1}\right)=\iota,$$

where E_t is the expectation operator conditional on information in period t and ι is an $n\times 1$ vector of 1s. This is the standard Euler condition, which equates the expected marginal cost and marginal benefit of delaying consumption one period. For example, in the case of time-separable preferences, m_{t+1} is the ratio of the marginal utility of future consumption to the marginal utility of current consumption. The unconditional version of the Euler equation is then

$$E(Rm)=\iota. \qquad (1)$$

Note that if there is a risk-free asset, then its gross return is $\frac{1}{E(m)}$. In the absence of a risk-free asset, we cannot pin down the mean of the SDF using return data.

Suppose we compute the least-squares projection of the SDF onto the linear space spanned by a constant and contemporaneous returns. The projection is of the form

(2) $$m = m_\nu + \varepsilon,$$

where

(3) $$m_\nu = \nu + \left(R - E(R)\right)' \beta,$$

$\beta \in \Re^n$, $\nu = E(m) = E(m_\nu)$, and ε is orthogonal to the constant as well as contemporaneous returns. This implies $E(\varepsilon) = 0$ and $E(R\varepsilon) = 0$. Together with the Euler equation (1), this implies $E(Rm) = E(Rm_\nu) = \iota$. Then

$$\mathrm{var}(m) = \mathrm{var}(m_\nu) + \mathrm{var}(\varepsilon) + 2\mathrm{cov}(m_\nu, \varepsilon).$$

By construction of m_ν, the projection error ε is orthogonal to m_ν, so $E(m_\nu \varepsilon) = 0$. Thus,

$$\mathrm{var}(m) = \mathrm{var}(m_\nu) + \mathrm{var}(\varepsilon) \geq \mathrm{var}(m_\nu),$$

meaning that a lower bound on the variance of a model's SDF m is the variance of m_ν. To find this lower bound, we need to know $\mathrm{var}(m_\nu)$.

From (3) it is easy to see that var $(m_\nu) = \beta' \Omega \beta$, where Ω is the variance-covariance matrix of asset returns. Since (2) and (3) describe a linear least-squares projection, we can estimate the projection coefficient β via OLS as $\beta = \Omega^{-1}\mathrm{cov}(R,m)$. Rewriting $\mathrm{cov}(R,m)$, we have $\beta = \Omega^{-1}(E(Rm) - E(m)E(R))$. Since the model implies $E(Rm) = \iota$, we can solve for β with

$$\beta = \Omega^{-1}\left(\iota - E(m)E(R)\right).$$

Thus, we can write $\mathrm{var}(m_\nu) = (\iota - E(m)E(R))'\Omega^{-1}(\iota - E(m)E(R))$. In terms of standard deviations, we can write the lower bound as

(4) $$\mathrm{std}(m) \geq \left\{ \left(\iota - E(m)E(R)\right)' \Omega^{-1} \left(\iota - E(m)E(R)\right) \right\}^{\frac{1}{2}}.$$

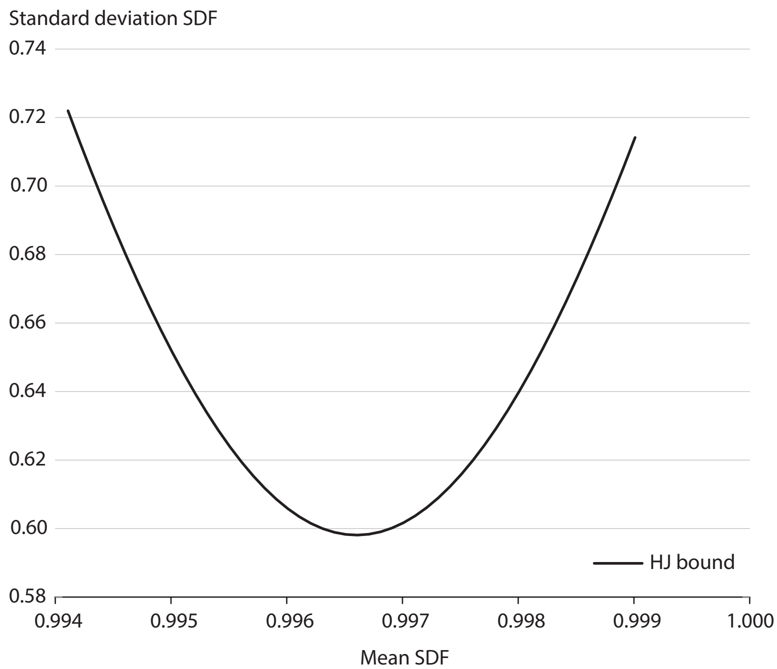
The right-hand side is the HJ bound. Note that the lower bound on the standard deviation of a model's SDF is a function of the mean of the model's SDF; so, it would seem like the lower bound depends on the model. However, we can generate a lower-bound *frontier* by picking different means. It is easy to see that the bound is a quadratic function of the mean SDF. A necessary condition for an SDF with mean $E(m)$ to be consistent with asset-return data is that it satisfies the inequality (4).

Computing the HJ bound frontier is straightforward. First, we calculate the sample mean of gross returns to use as a proxy for $E(R)$. Second, we calculate the variance-covariance matrix of the gross returns. Third, we choose a set of values for $E(m)$. For each value we compute the right-hand side of (4) to trace out a bound frontier.

Figure 1 illustrates the HJ bound using two asset returns from 1959:Q2 to 2019:Q2: the return on a 3-month Treasury bill and the return on the S&P 500. Both returns are transformed into real returns using the price deflator for personal consumption expenditures. (We use this deflator because when we conduct model evaluations later, we will be using personal

Figure 1

HJ Bound Frontier



NOTE: The figure depicts the set of admissible SDFs in mean-standard deviation space implied by stock and bond returns from 1959:Q2-2019:Q2.

consumption data.) The horizontal axis is $E(m)$, and the vertical axis is the HJ bound. The frontier is U-shaped, implying that SDFs with means far from the one associated with the least volatility will need to have higher volatility to satisfy the bound.[1]

3 EQUITY PREMIUM PUZZLE

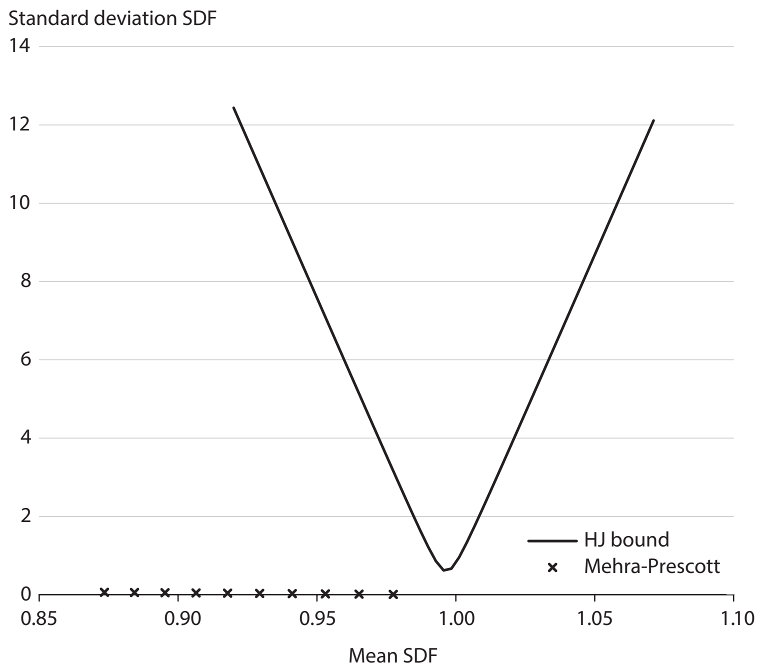
In this section, we use the HJ bound to illustrate the equity premium puzzle. To "test" a model using the HJ bound, we need the SDF implied by the model. As an example, suppose we want to check whether the Mehra and Prescott (1985) model is consistent with asset-return data. The preferences in their model are described by

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma}, \quad \sigma > 0,$$

where E_0 is the conditional expectation given information at time 0, c_t is the representative agent's consumption at time t, $\beta \in (0,1)$ is the subjective discount factor, and σ is the coefficient of relative risk aversion. (The preferences are assumed to be logarithmic when $\sigma = 1$.) The SDF for these preferences is given by

Figure 2

Mehra and Prescott (1985)



NOTE: Sample: 1959:Q2-2019:Q2. The HJ bound frontier in this figure is the same as in Figure 1. It also has the means and standard deviations of the Mehra-Prescott SDF for the range of the risk-aversion parameter (σ), from 1 to 10.

$$m_{t+1} = \beta\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma}.$$

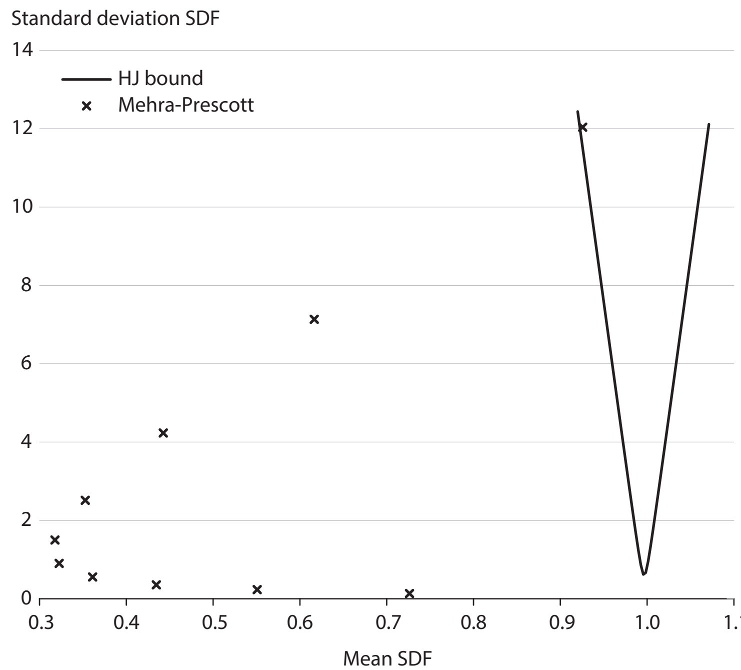
We can compute the time series of m_{t+1} using consumption data and parameter values for β and σ. Different values of β and σ imply different means, $E(m)$, and different volatilities, std(m). The question is whether there are any empirically plausible β and σ such that the pair ($E(m)$, std(m)) is inside the frontier. That is, given the mean of the model's SDF, the test is whether std(m) satisfies the bound in (4).

Figure 2 plots the same bound as in Figure 1 for the stock and bonds returns data. Figure 2 also plots the pairs ($E(m)$, std(m)) using quarterly nondurables and service consumption data from 1959:Q2 to 2019:Q2, for $\beta = 0.99$ and values of σ from 1 to 10. These values of β and σ are in the range investigated by Mehra and Prescott. The volatility for $\sigma = 1$ is the right most "x." As risk aversion is increased, the x's move to the left, but the increases in volatility are small.

For no value of σ is the bound satisfied. In fact, the volatilities of the SDF are far below the bound. We conclude that the model with this parameterization is rejected. A natural question is whether or not there exists a parameterization of the model that satisfies the bound.

Figure 3

High Risk Aversion in Mehra and Prescott (1985)



NOTE: Sample: 1959:Q2-2019:Q2. The figure illustrates the means and standard deviations of the Mehra-Prescott SDF for a wider range of the risk-aversion parameter (σ), relative to Figure 2, from 1 to 460.

To do so we increase risk aversion and find that the Mehra-Prescott model generates enough volatility to satisfy the bound when $\sigma = 460$; see Figure 3.

The high value of σ is unreasonable for two reasons. First, it implies an extreme aversion to risk. Second, it implies a high risk-free rate of 36 percent annually.[2] Figure 3 thus demonstrates the risk-free rate puzzle as well: The level of risk aversion that matches the observed equity premium comes at the cost of unreasonable values for the risk-free rate.

In sum, the Mehra-Prescott model of asset pricing is rejected for reasonable parameterizations of risk aversion. One needs implausibly high values of risk aversion to generate sufficient volatility to satisfy the bound. Moving forward, we need to find an SDF that generates higher volatility without high risk aversion.

Note that the above evaluation of the Mehra-Prescott model did not require us to solve the model or compute *equilibrium* asset returns. The test involved merely checking whether a necessary implication of the model was satisfied. We learned the same lessons that emerge from a full solution of the model. Another alternative to testing models using just the first-order conditions would be to estimate the Euler equation via GMM (generalized method of moments) as in Hansen and Singleton (1982) and then apply a J-test to the overidentifying restrictions. As is well known, this would lead to a statistical rejection of the Mehra-Prescott model. It would not, however, provide any guidance as to why the model was rejected and what to do to fix the model.

4 RESOLUTIONS

A similar procedure can be applied to two other popular asset-pricing models. Both are based on relaxing separability in the utility function: in one case state separability and in the other case time separability. Both add just one parameter to the Mehra-Prescott model, and both increase the volatility of the SDF.

Epstein and Zin (1991) and Weil (1989) generalize the time-separable preferences to allow for an independent parameterization of attitudes toward risk and intertemporal substitution. Following Weil (1989), these state-nonseparable preferences have a recursive representation:

$$V_t = U\left[c_t, E_t V_{t+1}\right],$$

where V is a von-Neumann-Morgenstern utility index and

$$U[c,V] = \frac{\left\{(1-\beta)c^{1-\rho} + \beta\left[1+(1-\beta)(1-\sigma)V\right]^{\left(\frac{1-\rho}{1-\sigma}\right)}\right\}^{\left(\frac{1-\sigma}{1-\rho}\right)} - 1}{(1-\beta)(1-\sigma)}.$$

The elasticity of intertemporal substitution is $1/\rho$, and σ is the coefficient of relative risk aversion. As shown by Weil (1989), the SDF for these preferences simplifies to

$$\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma}\right]^{\left(\frac{1-\sigma}{1-\rho}\right)}\left[R_{t+1}\right]^{\left(\frac{1-\sigma}{1-\rho}\right)-1},$$

where R_{t+1} is the return on the market portfolio.

Constantinides (1990) models consumers as habitual, in that levels of consumption in adjacent periods are complementary. That is, the time-nonseparable preferences of consumers (in a discrete-time, one-lag version of Constantinides, 1990) are given by

$$U_0 = E_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[c_t - \delta c_{t-1}\right]^{1-\sigma}}{1-\sigma},$$

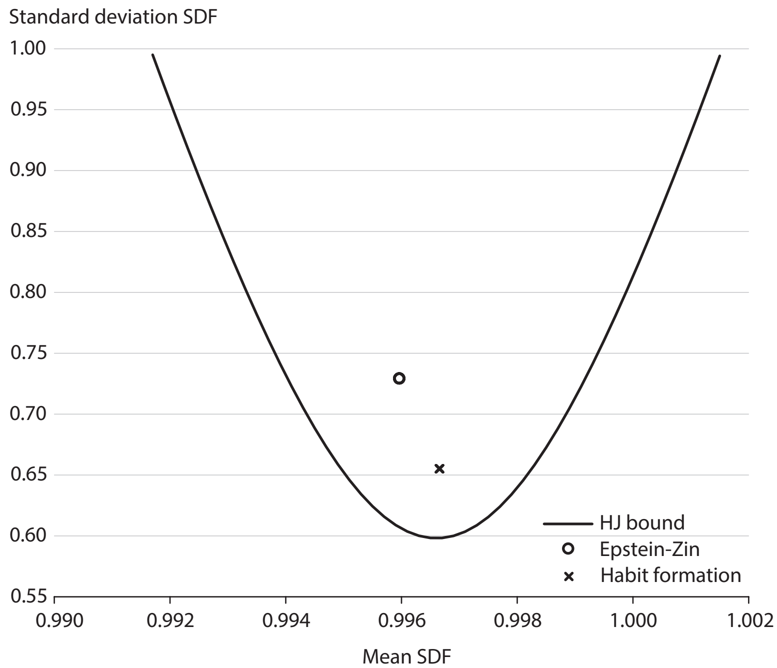
where $\delta > 0$. The representative agent's SDF is given by

$$m_{t+1} = \beta \frac{\left(c_{t+1} - \delta c_t\right)^{-\sigma} + \beta\delta E_{t+1}\left(c_{t+2} - \delta c_{t+1}\right)^{-\sigma}}{\left(c_t - \delta c_{t-1}\right)^{-\sigma} + \beta\delta E_t\left(c_{t+1} - \delta c_t\right)^{-\sigma}}.$$

Figure 4 plots the bound and the SDF volatilities for the two models. Again, the HJ bound is the solid curve, as in Figure 1; the "x" represents the habit SDF, while the "o" the Epstein-Zin SDF. For state-nonseparable preferences, the parameters are $\beta = 0.99$, $\rho = 0.9$, and $\sigma = 1.7$. For time-nonseparable preferences, the parameters are $\beta = 0.99$, $\delta = 0.8$, and $\sigma = 1.61$.[3] Both

Figure 4

Epstein-Zin and Habit-Formation Models



NOTE: Sample: 1959:Q2-2019:Q2. The figure illustrates the means and standard deviations of the SDFs for the Epstein-Zin and habit-formation preferences that satisfy the HJ bound. The HJ bound is the same as in Figure 1.

models satisfy the HJ bound by increasing the volatility of the SDF. In the case of the Epstein-Zin model, wealth, which is volatile, is part of the SDF. In the case of the time-nonseparable models, consumption growth is operated on by a difference operator, which increases volatility when raised to moderate powers.

4.1 Are the Resolutions Robust?

Our model evaluation shows that both Epstein-Zin and habit-formation models satisfy the HJ bound for apparently reasonable parameter values.[4] The evaluation was simple: It compared just two points—the volatility of the SDF and the HJ bound at the mean of the same SDF. The evaluation does not account for sampling variability in (i) asset-return data and (ii) consumption data. The sampling variability in (i) and (ii) might affect our inference on the model since the HJ bound is affected by (i) and the SDF is affected by (ii).

We now conduct a bootstrap experiment to take into account the two sampling variabilities and check whether the resolutions are robust to changes in the data sample. We adopt a variant of the bootstrap procedure in Otrok, Ravikumar, and Whiteman (2004): They first find parameters of the asset-pricing models that satisfy the HJ bound for the whole post-WWII sample. Using these parameters, they then show that the models do not satisfy the bound for subsamples.

Our bootstrap experiment here, however, investigates how the HJ bound and the volatility of the SDF vary across artificial samples drawn from the full data set (1959:Q2-2019:Q2). To do this we compute the time series for the representative agent's SDF for the two successful asset-pricing models in Section 3 using consumption growth data. We then use a bootstrap procedure to sample a vector of asset returns and the SDF. We bootstrap the entire vector—consumption, equity return, and bond return—so that the observed correlation properties between the two returns and the SDF are maintained in our experiment.

The bootstrap procedure is as follows:

(i) Use the parameters from Section 3 and observed consumption growth data to get time series for the SDFs of the models. (The parameters are $\beta = 0.99$, $\rho = 0.9$, and $\sigma = 1.7$ for the Epstein-Zin SDF and $\beta = 0.99$, $\delta = 0.8$, and $\sigma = 1.61$ for the habit SDF.)
(ii) Draw (with replacement) a time series of length 241 from the joint "empirical" distribution of the SDFs, equity returns, and T-bill returns. That is, for each period we draw a 3-tuple (SDF, R^{equity}, $R^{\text{T-bill}}$).
(iii) Calculate the mean and volatility of the SDF.
(iv) Calculate the HJ bound using the time series for equity and T-bill returns at the mean SDF.
(v) Repeat steps ii-iv 1,000 times.

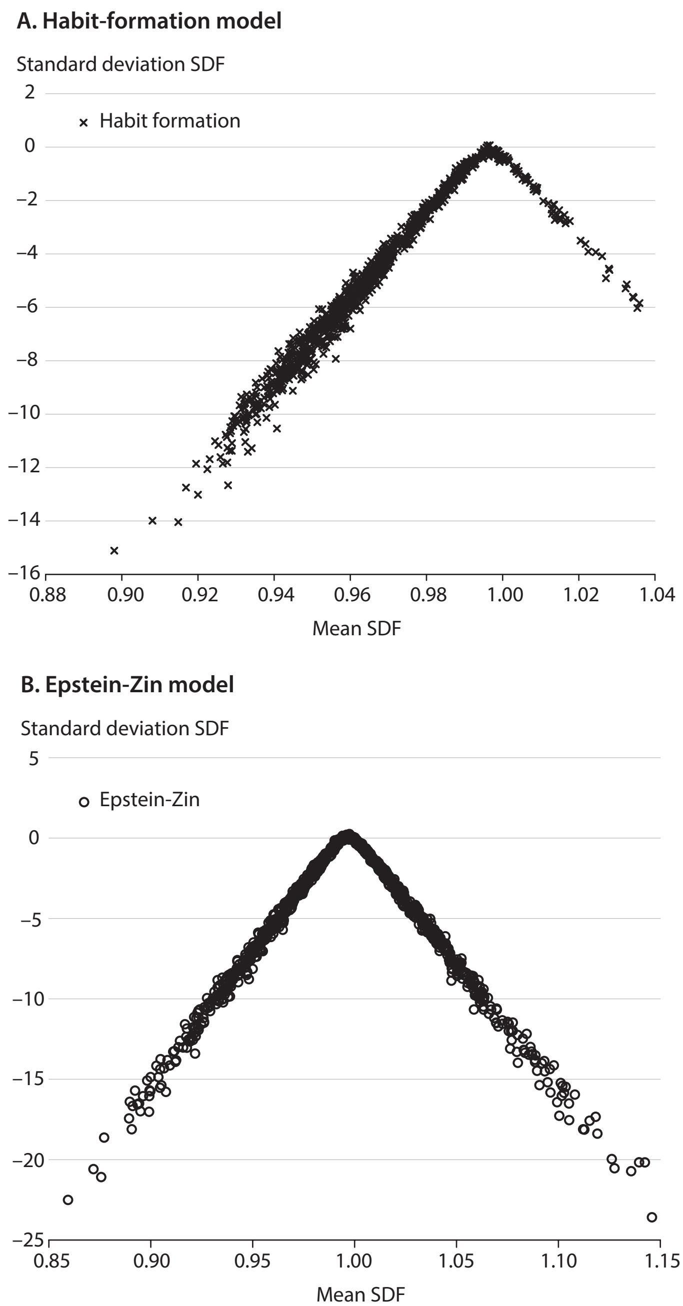
Figure 5 is a scatter plot of the distance between the HJ bound and SDF volatility, calculated as the SDF volatility minus the HJ bound, for each of the 1,000 bootstrap simulations. Panel A plots the habit-formation model, while Panel B plots the Epstein-Zin model. The striking feature of these figures is that the distance is almost always negative, implying that the models miss the bound in most simulations. In fact, the habit model misses in 96 percent of the simulations and the Epstein-Zin model misses in 95 percent of the simulations.

Burnside (1994) casts the distance to the bound in a GMM framework. He studies statistical measures of the distance between the HJ bound and SDF volatility in the *time-separable* model and argues that the over rejection is partly due to variations in the mean of the SDF. Cecchetti, Lam, and Mark (1994) also show that in the context of models with time-separable preferences and habit formation preferences, much of the variability in the distance is due to the uncertainty in estimating the mean of the SDF. To their point, even if we consider only the lowest possible bound from the bootstrap simulations, we will reject the models most of the time. The reason is that the mean of the SDF varies greatly across bootstrap samples. Since the bound itself rises rapidly for $E(m)$ different from 0.99, the distance to the bound becomes large and leads to a rejection.[5]

Statistically speaking, for the HJ bound to be a useful evaluation device, the test should not reject a true model. Specifically, suppose one uses observed consumption data to solve an asset-pricing model, that is, compute the *equilibrium* asset returns implied by the model. Then the test based on the distance between the HJ bound associated with the equilibrium returns and the volatility of the model SDF should not reject the true model. One can judge the test by simulating the true model many times and counting the number of times the the HJ bound is violated. Gregory and Smith (1992) conduct this exercise for time-separable preferences and conclude that the true model is rejected frequently.

Figure 5

Bootstrap Simulations



NOTE: Sample: 1959:Q2-2019:Q2. The figure plots the distance between the standard deviation of the SDF and the HJ bound evaluated at the mean of the SDF for each draw of the bootstrap. The parameters for the Epstein-Zin SDF are $\beta = 0.99$, $\rho = 0.9$, and $\sigma = 1.7$; the parameters for the habit SDF are $\beta = 0.99$, $\delta = 0.8$, and $\sigma = 1.61$.

A formal statistical evaluation involves calculating rejection rates based on critical values of the test statistic, as in Burnside (1994) and Cecchetti, Lam, and Mark (1994). However, Otrok, Ravikumar, and Whiteman (2002) show that tests based on the distance to the HJ bound are non-pivotal in finite samples: The finite-sample critical values depend upon the SDF parameters: risk aversion and the discount factor. Therefore, one has to calculate parameter-specific critical values for each point in the null hypothesis of interest. Nevertheless, for the case of time-separable preferences, Otrok, Ravikumar, and Whiteman (2002) show that the finite-sample distribution of the test statistic associated with the risk-neutral case is extreme. The critical values for the risk-neutral case deliver type-I errors no larger than intended, regardless of risk aversion or the discount factor. They also show that the maximal type-I error critical values for time-separable preferences are appropriate for habit formation as well as state nonseparable preferences. Their conclusion is that the HJ bound is indeed a useful statistical evaluation device, in that type-I errors can be controlled, while type-II error rates are acceptably small. Using their finite-sample critical values, they report evidence against time-separable preferences and mixed evidence for Epstein-Zin and habit preferences.

5 ASSET-PRICING IMPLICATIONS OF BUSINESS CYCLE MODELS

Our focus so far in this article has been on asset-pricing models and financial returns data typically used in the asset-pricing literature. The HJ bound is also useful for analyzing the asset-pricing implications of business cycle models. Such an approach is useful since a business cycle model is typically solved with a first-order approximation, which eliminates risk premia. Higher-order solutions are possible but costly for moderate-sized models. An early approach to using the HJ bound in the context of a business cycle model was Tallarini (2000). That paper first showed that risk-aversion per se did not affect the business cycle behavior of standard macroeconomic aggregates. It then showed that the SDF from that model did satisfy the HJ bound with sufficiently high risk aversion.

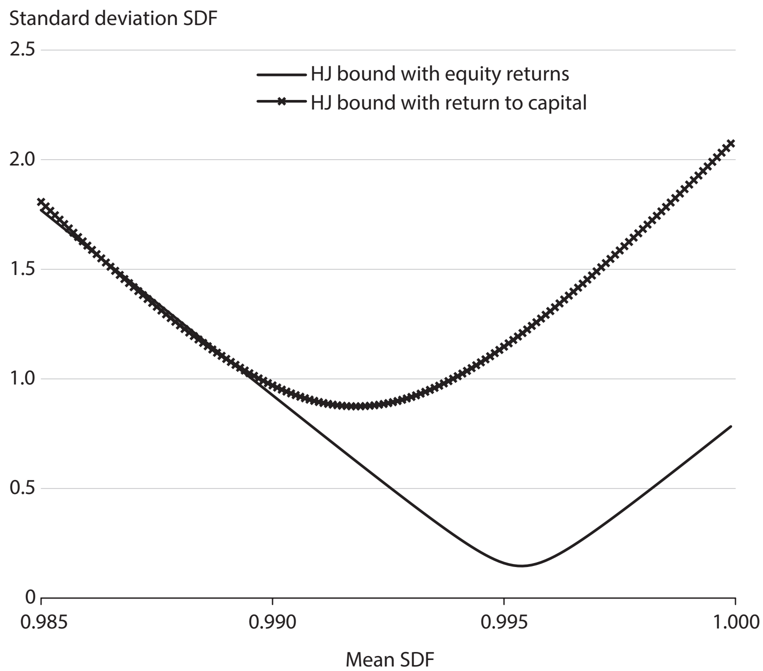
Typically, in business cycle models that study asset-pricing implications, asset return is measured by the S&P 500. Gomme, Ravikumar, and Rupert (2011) argue that business cycle theory does not necessarily imply using financial return to measure the return to capital. They construct the return to capital in the United States using NIPA statistics on capital income and capital stock. They show that while the mean return is roughly the same for equity returns and the return to capital, the NIPA return to capital is less volatile.

In Figure 6 we construct the HJ bound using the return to capital as in Gomme, Ravikumar, and Rupert (2011) and the return to the 3-month Treasury bill. We also plot the HJ bound from Figure 1 for comparison. The sample period here is 1959:Q2-2008:Q4. The bound for the return to capital is significantly higher. This follows from the fact that the return to capital is less volatile than the return to equity, which leads to a sharper set of restrictions on the set of admissible SDFs.

The implication for business cycle models is that the asset-pricing puzzle is in fact more challenging than the one we see with financial return data. As in Figure 2, the time-separable model does not generate enough volatility to satisfy the HJ bound with equity returns, so it

Figure 6

Return to Capital: Gomme, Ravikumar, and Rupert (2011)



NOTE: Sample: 1959:Q2-2008:Q4. The figure depicts the set of admissible SDFs in the mean-standard deviation space implied by stock and bond returns as in Figure 1. It also depicts the set of admissible SDFs implied by return to capital from Gomme, Ravikumar, and Rupert (2011) and bond return.

certainly will not satisfy the bound with return to capital. While it may be possible to find time- or state-nonseparable preferences that satisfy this bound, they will still suffer from the stability problem we documented earlier with these resolutions. In addition, in the case of time-non-separable preferences, Otrok (2001) shows that the data prefer only moderate amounts of habit formation, which will not generate much volatility in the SDF.

6 CONCLUSION

The Hansen-Jagannathan bound is a helpful tool for understanding asset-pricing implications. By characterizing the admissible set of SDFs, we can use the bound to test proposed SDFs. Further, we can understand what types of asset-return data will pose greater difficulty for an asset-pricing model. Lastly, the bound can be constructed with only the first and second moments of asset-return data, hence implementation of the bound requires no computing power beyond a spreadsheet program.

The HJ bound uses the means, variances, and contemporaneous correlations of asset-return data to construct the lower bound that an SDF must satisfy. Otrok, Ravikumar, and Whiteman (2007) develop a volatility bound that uses serial correlation properties of the return data as well. This generalization allows for an evaluation of whether models fail to match the data in the long run, at business cycle frequencies, etc. That is, the generalized bound can help identify the frequencies at which a model violates the necessary conditions. A business cycle model that violates the bound at business cycle frequencies might be unacceptable, but violations at other frequencies might not be a cause for concern. The generalization involves projecting the SDF onto the space of current, past, and future returns. Because the projection is onto a larger space than that for the HJ bound, the generalized bound is tighter than the HJ bound. They find that the state-nonseparable SDF satisfies the bound at business cycle frequencies, while the time-nonseparable SDF does poorly at those frequencies. Open questions for future research are whether the resolutions at those frequencies are stable and whether these SDFs can satisfy the generalized bound at business cycle frequencies when the bound is constructed with Gomme, Ravikumar, and Rupert (2011) capital-return data. ■

APPENDIX: HJ BOUND AND THE SHARPE RATIO

The Sharpe ratio is the mean excess return on an asset (relative to the risk-free rate) divided by the standard deviation of that asset's return. The Sharpe ratio measures how the market views risk: A higher Sharpe ratio implies that the market demands a higher return for a given level of risk. The connection between the HJ bound and the Sharpe ratio presented here follows Cochrane (2001).

Consider the unconditional Euler equation used to price assets:

$$\iota = E\left(R^{\text{equity}} m\right).$$

For the sake of exposition we will assume that the only asset is equity, with return R^{equity}, and there is a risk-free rate such that $R^f = \frac{1}{E(m)}$. We can write the right-hand side of the equation as

$$E\left(R^{\text{equity}} m\right) = E(m) E\left(R^{\text{equity}}\right) + \rho_{R^{\text{equity}},m} \sigma_{R^{\text{equity}}} \sigma_m,$$

where $\rho_{R^{\text{equity}},m}$ is the correlation of equity returns with m and σs represent standard deviations. Next, divide through by $E(m)$ to get

$$\frac{1}{E(m)} = E\left(R^{\text{equity}}\right) + \frac{\rho_{R^{\text{equity}},m} \sigma_{R^{\text{equity}}} \sigma_m}{E(m)}.$$

Replacing $\frac{1}{E(m)}$ with R^f, dividing by $\sigma_{R^{\text{equity}}}$, and rearranging terms yields

$$\frac{E\left(R^{\text{equity}}\right) - R^f}{\sigma_{R^{\text{equity}}}} = -\frac{\rho_{R^{\text{equity}},m} \sigma_m}{E(m)}.$$

Since $-1 \le \rho_{R^{\text{equity}},m} \le 1$, we have the inequality

$$\frac{E\left(R^{\text{equity}}\right) - r^f}{\sigma_{R^{\text{equity}}}} \le \frac{\sigma_m}{E(m)}.$$

The left-hand side is the Sharpe ratio. For given $E(m)$, a higher Sharpe ratio implies that the lower bound on SDF volatility is higher.

NOTES

1 Matlab code and data used for this and all subsequent examples in this article can be found on Christopher Otrok's REPEC webpage: https://ideas.repec.org/e/pot2.html.

2 For the value of σ that satisfies the bound, $Em = 0.9258$, or a quarterly return of $1/Em = 1.0801$.

3 Note that σ is not the coefficient of risk aversion in the habit model, though it is proportional to various measures of risk aversion. See Boldrin, Christiano, and Fisher (1997).

4 Since we can dismiss the Mehra-Prescott model for achieving the bound with only unreasonable amounts of risk aversion, we will focus on only these two models in this section.

5 The value 0.99 is in the lower part of the bound in Figure 1.

REFERENCES

Boldrin, M.; Christiano, L.J. and Fisher, J.D.M. "Habit Persistence and Asset Returns in an Exchange Economy." *Macroeconomic Dynamics*, 1997, *1*, pp. 312-32; https://doi.org/10.1017/S1365100597003027.

Burnside, C. "Hansen-Jagannathan Bounds as Classical Tests of Asset Pricing Models." *Journal of Business and Economic Statistics*, 1994, *12*, pp. 57-79; https://doi.org/10.1080/07350015.1994.10509991.

Cecchetti, S.G.; Lam, P-S. and Mark, N.C. "Testing Volatility Restrictions on Intertemporal Marginal Rates of Substitution Implied by Euler Equations and Asset Returns." *Journal of Finance*, 1994, *49*, pp. 123-52; https://doi.org/10.1111/j.1540-6261.1994.tb04423.x.

Cochrane, J.H. *Asset Pricing*. Princeton University Press, 2001.

Constantinides, G.C. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, 1990, *98*, pp. 519-43; https://doi.org/10.1086/261693.

Epstein, L.G. and Zin, S.E. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis." *Journal of Political Economy*, 1991, *99*(2), pp. 263-86; https://doi.org/10.1086/261750.

Gomme, P.; Ravikumar, B. and Rupert, P. "The Return to Capital and the Business Cycle." *Review of Economic Dynamics*, 2011, *14*(2), pp. 262-78; https://doi.org/10.1016/j.red.2010.11.004.

Gregory, A.W. and Smith, G.W. "Sampling Variability in Hansen-Jagannathan Bounds." *Economics Letters*, *38*, pp. 263-67; https://doi.org/10.1016/0165-1765(92)90068-A.

Hansen, L.P. and Jagannathan, R. "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy*, 1991, *99*, pp. 225-62; https://doi.org/10.1086/261749.

Hansen, L.P. and Singleton, K.J. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica*, 1982, *50*(5), pp. 1269-86; https://doi.org/10.2307/1911873.

Ljungqvist, L. and Sargent, T.J. *Recursive Macroeconomic Theory*. MIT Press, 2018.

Mehra, R. and Prescott, E.C. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 1985, *15*, pp. 145-61; https://doi.org/10.1016/0304-3932(85)90061-3.

Otrok, C.; Ravikumar, B. and Whiteman, C.H. "Evaluating Asset-Pricing Models Using the Hansen-Jagannathan Bound: A Monte Carlo Investigation." *Journal of Applied Econometrics*, 2002, *17*, pp. 149-174; https://doi.org/10.1002/jae.640.

Otrok, C.; Ravikumar, B. and Whiteman, C.H. "Stochastic Discount Factor Models and the Equity Premium Puzzle." Unpublished manuscript, 2004; https://ideas.repec.org/p/pra/mprapa/22938.html.

Otrok, C.; Ravikumar, B. and Whiteman, C.H. "A Generalized Volatility Bound for Dynamic Economies." *Journal of Monetary Economics*, 2007, *54*(8), pp. 2269-90; https://doi.org/10.1016/j.jmoneco.2007.06.028.

Otrok, C. "On Measuring the Welfare Cost of Business Cycles." *Journal of Monetary Economics*, 2001, *47*(1), pp. 61-92; https://doi.org/10.1016/S0304-3932(00)00052-0.

Tallarini, T. "Risk-Sensitive Real Business Cycles." *Journal of Monetary Economics*, 2000, *45*(3), pp. 507-32; https://doi.org/10.1016/S0304-3932(00)00012-X.

Weil, P. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle." *Journal of Monetary Economics*, 1989, *24*, pp. 401-21; https://doi.org/10.1016/0304-3932(89)90028-7.

Reconstructing the Great Recession

Michele Boldrin, Carlos Garriga, Adrian Peralta-Alva, and Juan M. Sánchez

This article uses dynamic equilibrium input-output models to evaluate the contribution of the construction sector to the Great Recession and the expansion preceeding it. Through production interlinkages and demand complementarities, shifts in housing demand can propagate to other economic sectors and generate a large and sustained aggregate cycle. According to our model, the housing boom (2002-07) fueled more than 60 percent and 25 percent of employment and GDP growth, respectively. The decline in the construction sector (2007-10) generates a drop in total employment and output about half of that observed in the data. In sharp contrast, ignoring interlinkages or demand complementarities eliminates the contribution of the construction sector. (JEL E22, E32, O41)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2020, *102*(3), pp. 271-311.
https://doi.org/10.20955/r.102.271-311

1 INTRODUCTION

With the onset of the Great Recession, U.S. employment and gross domestic product (GDP) fell dramatically and then took a long time to return to their historical trends. There is still no consensus about what exactly made the recession so deep and the subsequent recovery so slow. In this article we evaluate the role played by the construction sector in driving the boom and bust of the U.S. economy during 2001-13.

The construction sector represents around 5 percent of total employment, and its share of GDP is about 4.5 percent. Mechanically, the macroeconomic impact of a shock to the construction sector should be limited by these figures; we claim it is not. Rather, we claim one reason why the Great Recession was particularly deep and persistent is that the construction sector and housing consumption are strongly interconnected to the rest of the economy.

Michele Boldrin is professor of economics at Washington University in St. Louis. Carlos Garriga is a vice president and economist at the Federal Reserve Bank of St. Louis. Adrian Peralta-Alva is a senior economist at the International Monetary Fund. Juan M. Sánchez is an assistant vice president and economist at the Federal Reserve Bank of St. Louis.

The authors are grateful for the stimulating discussions with Morris Davis; Bill Dupor; Ayse Imrohoroglu; Rody Manuelli; Alex Monge-Naranjo; Alp Simsek; Jorge Roldos; Paul Willen; and the seminar participants at the Bank of Canada, Southern Methodist University, International Monetary Fund, 2012 Wien Macroeconomic Workshop, 2012 ITAM Summer Camp in Macroeconomics, 2012 SED, 2015 GBUREES, Stony Brook, and Queens College CUNY. Constanza Liborio, James D. Eubanks, and You Chien provided excellent research assistance.

Such linkages are important at the production stage (purchases of intermediate goods) but also at the final consumption stage (broad demand complementarities).

Our vision of how a shock to the demand for housing travels to the rest of the economy is the following. In response to a demand-driven housing boom, the construction sector leads the rest of the economy by fueling an expansion through its purchases of inputs. The expansion generates an increase in consumption (housing and nonhousing) as well as investment (residential and capital). This continues either until a new steady state is reached or, as in the historical case we study, a sudden drop in housing demand generates a decline in construction output. This translates into a general reduction of demand (for intermediate inputs and complementary consumption goods), thereby propagating and magnifying, again, the negative sectoral shock. Further, a sudden drop in the demand for housing also generates a slow recovery because the excessive inventory of housing units takes a long time to be absorbed, hence the particularly long delay in the aggregate recovery.

In the empirical analysis, we construct measures of sectoral interlinkages (multipliers) and show that the construction sector is one of the most interconnected in the economy. We use this to quantify the contribution of the construction sector during the period 2002-13 and estimate it to have been unusually large. Construction is capable of accounting for about 52 percent of the decline in total employment and 35 percent of the decline in aggregate gross output. Eliminating the production multipliers weakens the impact on total employment to 20.8 percent and on gross output to 19.3 percent.

In a simplified version of the model, we illustrate the importance of production interlinkages and demand complementarities. This exercise provides a set of sufficient conditions under which the presence of interlinkages generates larger effects in aggregate employment and output than in their absence. The algebra indicates that, for the amplification effect to exist, the sectoral interlinkages must be asymmetric, with the construction sector buying relatively more inputs from the rest of the economy than vice versa. This condition is supported—by more than two orders of magnitude—by estimates from the U.S. input-output table from the Bureau of Economic Analysis (BEA). To generate a multiplier effect via interlinkages, it is also necessary to have an elasticity of substitution between housing and consumption goods lower than 1.[1] With an elasticity of substitution larger than 1, a decline in housing demand generates the reallocation of productive inputs away from the construction sector and a boom in the nonhousing sectors, which may (more than) compensate for the decline due to the supply-side interlinkages. With a unitary elasticity, these two effects cancel out.

One of the limitations of the static model is that it does not allow study of the dynamic adjustment of consumption, residential investment, and productive capital. It also ignores the process of the adjustment of relative prices, and it is not ideal for quantitative purposes. To overcome these limitations, we solve the full dynamic general-equilibrium model numerically. We use that model to answer the following question: If demand for housing shifts exogenously over time to match the observed dynamics of employment in the construction setor, what will happen to the remaining macro quantities and prices?[2]

Our simulations indicate that, in the presence of sectoral interlinkages and consumption complementarities, the size of the boom-bust cycle in total employment and output is substantially larger and more coherent with historical observations than otherwise. During the demand-driven housing boom, both sectors in our model expand and contribute to the growth of output and employment, by 2 percent and 2.5 percent, respectively. During the housing bust, the decline in output is 3.3 percent and in total employment is 3.8 percent. The model also captures the leading role of construction during booms and busts (see Leamer, 2007) and the comovement with nonhousing expenditures and investment. Further, the separation of productive capital and residential structures, together with the irreversibility constraints, introduces an asymmetry between expansions and recessions similar to that observed by many previous researchers and that is traditionally hard to obtain in most real business cycle models.[3] As in the theoretical exercise, the dynamic quantitative simulations show that reducing the importance of either the sectoral interlinkages or the demand complementarities weakens the transmission mechanism. Moreover, the model without linkages also fails to capture the lead-lag pattern of housing and consumption expenditures observed in the data. These results indicate that modeling production linkages provides a quantitatively relevant transmission channel.

The burst of the real estate "bubble" might have substantially lowered potential output and created a substantial "displacement effect," for both labor and capital, which took quite some time to absorb. Some researchers have referred to this displacement effect as a worsening of labor frictions. For example, Arellano, Bai, and Kehoe (2019) and Ohanian and Raffo (2012) attribute the Great Recession primarily to this factor. Since our model captures significant declines in employment and output in the absence of such frictions, we also perform a business cycle accounting exercise on simulated data from the model. Through the lens of the one-sector neoclassical growth model, the presence of intersectoral linkages, movements in relative prices, and shifts in housing demand can be interpreted as "distortions." Business cycle accounting would attribute the recession generated in the model to the labor wedge. In our model, the magnitude of the worsening of the labor wedge is about 62 percent of the total change observed in the data. Importantly, in both our model and the data, the worsening is due to the consumer side of the labor wedge and not to differences between wages and the marginal product of labor.

Obviously, the fluctuations of the construction sector cannot fully account for the dynamics of employment and output since 2002. Other relevant factors not incorporated into the analysis are important. Many suggest (Black, 1995; Hall, 2011; and Kocherlakota, 2012) that high interest rates could be responsible for the slow recovery. These authors argue that even in models with perfect competition and price flexibility (i.e., lacking the typical frictions of New-Keynesian business cycle models), too-high interest rates may result in substantially lower levels of output and employment. Since some interest rates appear to be currently constrained by the zero lower bound, such analyses appear particularly pertinent. Others argue that the level of uncertainty (Bloom, 2009, and Arellano, Bai, and Kehoe, 2019), government policies (Herkenhoff and Ohanian, 2011), and excessive debt overhang in the economy (Garriga, Manuelli, and Peralta-Alva, 2019; Herkenhoff and Ohanian, 2012; and Kehoe, Ruhl, and Steinberg, 2013) may be responsible for the lackluster recovery. Our exercise is silent with respect to these factors.

The remainder of the article is organized as follows. Section 2 connects our work with the related literature. Section 3 evaluates the importance of production multipliers first in the data, using the input-output tables of the U.S. economy, and then theoretically, using a stylized two-sector model. Section 4 presents the quantitative multisector model, illustrates the key mechanism at work in the analysis, and quantifies the importance of production linkages and demand complementarities. Section 6 compares the implications of the model in terms of business cycle accounting methodology, and Section 7 offers some concluding comments.

2 RELATED LITERATURE

There is a large literature studying the connection between housing and the macroeconomy,[4] for example, Gervais (2002), Campbell and Hercowitz (2005), Leamer (2007), Fisher (2007), and Davis and Van Nieuwerburgh (2015). Most of these papers explore the effects of housing at the traditional business cycle frequency, ignoring large swings in growth rates as in the case of the Great Recession and the precedent boom. For example, Davis and Heathcote (2005) study the comovement of residential and nonresidential investment in a dynamic multi-sector model with interlinkages and unitary elasticity of substitution between housing and other goods. Our results show that intersectorial production linkages have large propagation and magnification effects on most macroeconomic variables and that these effects are larger in the presence of demand complementarities.

A small strand of the growth-theoretical literature has long argued that asymmetries in the input-output structure of multisectoral neoclassical models amplify the effect of sectoral shocks and may even generate endogenous cycles for appropriate configurations of the parameters, for example, Benhabib and Nishimura (1979), Boldrin and Deneckere (1990), and Long and Plosser (1983). This theoretical theme has recently seen a revival in a growing literature, both theoretical and applied, stressing that the intersectoral composition of the production sector is an important source of propagation of idiosyncratic sectoral shocks (i.e., Horvath, 1998 and 2000; Carvalho, 2010; Foerster, Sarte, and Watson, 2011; Gabaix, 2011; Acemoglu et al., 2012; Carvalho and Gabaix, 2013; Caliendo et al., 2014; Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015; and Atalay, Dratzburg, and Wan, 2018).[5] This literature focuses on idiosyncratic technological shocks affecting the shape of the production possibility frontier and thereby generating movements of aggregate output. These sectoral supply shocks have been proved to be an important channel through which local perturbations may lead to aggregate fluctuations, absent any change in the composition of aggregate demand. For our part, we study how, absent any technological variation, sectoral demand shocks may also cause aggregate fluctuations. In the case of the Great Recession, it is hard to identify a specific "sectoral production possibility shifter," while there clearly was a dramatic drop in the demand for housing. Interestingly, our analysis shows that technological asymmetries—for example, that a one-sector model is a poor representation of the underlying production possibility set—play a crucial role also in the case of demand shocks. It also shows that demand complementarities, largely ignored in the business cycle literature, are in fact quantitatively relevant and should not be ignored.

Further along the sectoral-aggregate divide, Li and Martin (2014) study the transmission of shocks using dynamic factor methods and explicitly look at input-output linkages. They find that a significant part of traditionally defined aggregate fluctuations are driven by "sector-specific shocks." In the case of the Great Recession, more than half of aggregate volatility is accounted for by an additional aggregate shock—which they label the "wedge factor"—emerging only during this period. Most crucially, and consistently with our bottom line, they find that shocks originating in the construction sector generate the largest spillover effects over time, dominating those of all the other sectors. In our model, the driving force is an exogenous shifter in housing demand that acts like a wedge factor.

There is also an extensive literature that explores the role of financial conditions as drivers of the Great Recession and of the delayed recovery (i.e., Black, 1995, and Bloom, 2009; Christiano Motto, and Rostagno, 2010; Arellano, Bai and Kehoe, 2019; Gertler and Karadi, 2011; Hall, 2011; Jones, Midrigan, Philippon, 2018; Kocherlakota, 2012; Jermann and Quadrini, 2012; Brunnermeier and Sannikov, 2013; He and Krishnamurthy, 2019; and Mitman, Kaplan, and Violante, 2020). Most of the literature abstracts from the role of housing during this episode, with a few exceptions. Among them is Garriga, Manuelli, and Peralta-Alva (2019). In their model, an increase in the cost of housing financing generates a collapse of house prices, inducing a recession through deleveraging. Similarly, Martinez, Hatchondo, and Sánchez (2015) use heterogeneous agent models to analyze the aggregate effects of a house price decline and of its propagation to the rest of the economy through household balance sheets and housing defaults. Iacoviello and Pavan (2013) argue that a tightening of household budgets, due to the drop in real estate wealth, induced a sharp decline in aggregate consumption. Rognlie, Shleifer, and Simsek (2018) explore the aggregate effect of insufficient housing demand resulting from a period of overbuilding. Our article is complementary to this literature because, once again, we take as granted the drop in housing demand and then study its supply-side propagation due to sectoral interlinkages.

3 CONSTRUCTION IN AN INPUT-OUTPUT ECONOMY

This section first provides empirical evidence and then a simple theoretical evaluation of the importance of interlinkages. The data analysis places special attention on the Great Recession but also uses detailed U.S. sectoral input-output data for the period 1990-2013. The theoretical framework provides a set of sufficient conditions for the amplification mechanism to work.[6]

3.1 Construction and Aggregate Fluctuations: 1990-2013

For the analysis of economic fluctuations, it is common to use aggregate data for the whole postwar period. Unfortunately, the current availability of uniform input-output data is limited to the years 1990-2014. According to the BEA, the U.S. economy has experienced three recessions (1990-91, 2000-01, and 2007-09) during that interval of time. To evaluate the *direct contribution* of the construction sector to each episode, Table 1 summarizes the changes in employment and real income for the construction and private sectors. We measure the direct contribution of construction as the ratio of the change in construction to each total

Table 1

The Role of Construction in the Past Three Recessions

	Employment (millions)		Real income (billions of dollars)	
Recessions	Construction sector	Private sector	Construction sector	Private sector
1990-91				
Peak, 1990:Q3	5.24	109.6	346.0	7,455
Trough, 1991:Q1	4.93	108.7	320.6	7,400
Difference	–0.31	–0.87	–25.4	–54.9
Percent accounted for by construction	35.4		46.3	
2000-01				
Peak, 2000:Q1	6.84	132.6	610.4	10,870
Trough, 2001:Q4	6.79	131.0	598.7	10,629
Difference	–0.05	–1.55	–11.70	–241.1
Percent accounted for by construction	3.3		4.9	
2007-09				
Peak, 2007:Q4	7.53	137.9	682.3	12,586
Trough, 2009:Q2	6.09	131.0	540.5	11,852
Difference	–1.4	–6.9	–141.8	–734.0
Percent accounted for by construction	20.3		19.3	

SOURCE: BEA and BLS.

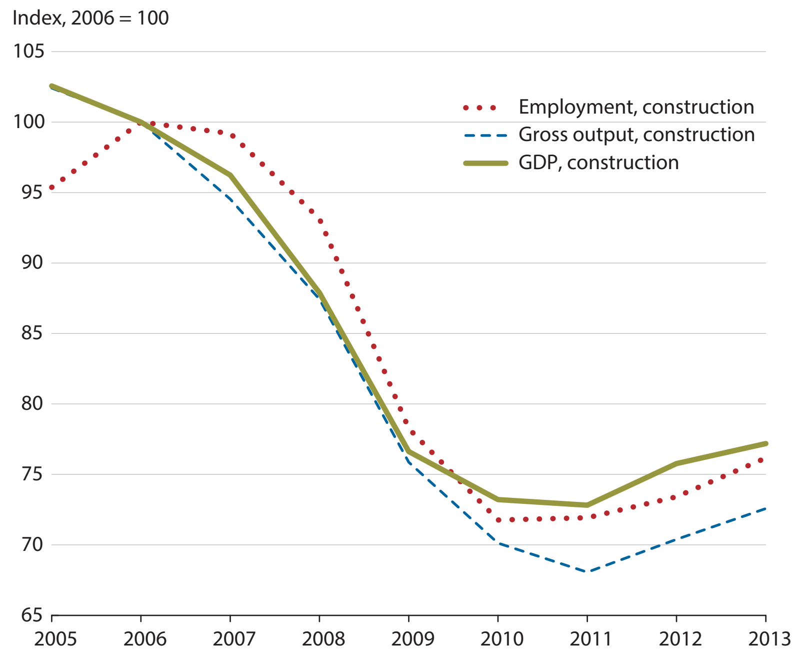
change. The top panel of Table 1 reminds us that the 1990-91 recession was mild: Between the peak of 1990:Q3 and the trough of 1991:Q1, for the private sector, employment and income each declined by less than 1 percent. In relative terms, the declines in the construction sector were sizeable: slightly less than 6 percent for employment and more than 7 percent for income. The middle panel of Table 1 shows the recession that started in 2000:Q1 and ended in 2001:Q4. This recession was slightly more severe: for the private sector, employment fell by more than 1 percent and income by more than 2 percent. However, the declines in the construction sector were almost negligible and the shares of the aggregate declines they accounted for were small.

The Great Recession started in 2007:Q4 and lasted until 2009:Q2. During this period, total employment decreased by roughly 7 million jobs. Table 1 shows that the direct contributions of the drops in construction employment and real income were 20.3 percent and 19.3 percent, respectively. This recession was dramatically bigger than the previous two, and the size of the drop in employment was 18.6 percent, not 5.9 percent as in 1990-91 or 0.7 percent as in 2000-01.

These calculations ignore the fact that construction leads the cycle and that during the Great Recession this industry went into recession 18 months before the overall economy. Measuring the decline from the perspective of the construction cycle shows that employment fell from 7.7 million (2006:Q3) to 5.5 million (2011:Q1) and recovered little thereafter. Figure 1

Figure 1

The Construction Sector During the Great Recession



SOURCE: BEA.

shows that employment, gross output, and GDP in the construction sector each dropped about 30 percent, whereas Table 1 infers employment dropped by 18.6 percent.

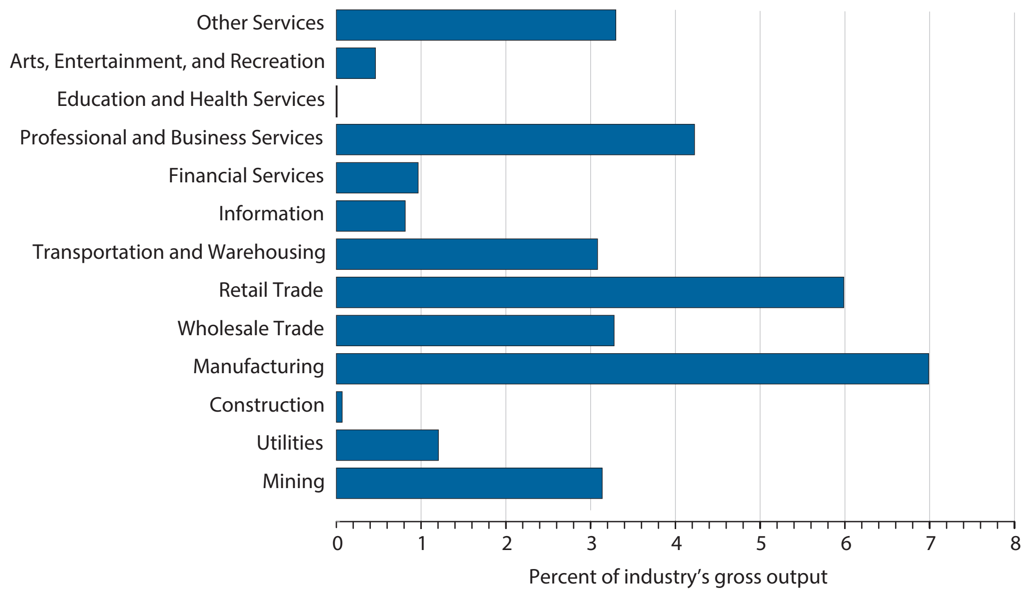
3.2 Evidence and Implications of Production Interlinkages

This subsection uses input-output data to provide a rough estimate of the role of interlinkages in the aggregate amplification of sectoral demand shocks during the cycle leading up to the Great Recession. Despite its relatively small size, the contribution of the construction sector to the Great Recession was a combination of two factors: the large size of the shocks affecting the sector and the sector's strong interlinkages with suppliers. One way of measuring sectoral interlinkages is through the purchases of a sector from other sectors, expressed as a percentage of the total output of the latter. For the period 1990-2013, these calculations are reported in Figure 2.

These numbers show by how much the total demand for the (gross) output of these sectors would immediately decrease if construction demand vanished. For instance, the total demand for the manufacturing sector would immediately decrease by 7 percent if the construction sector vanished. This measures only the *direct effect*: Because each sector purchases goods and services from other sectors as inputs, the process continues—virtually for an infinite number of steps—until it converges, thereby inducing a "production multiplier" effect. This multiplier

Figure 2

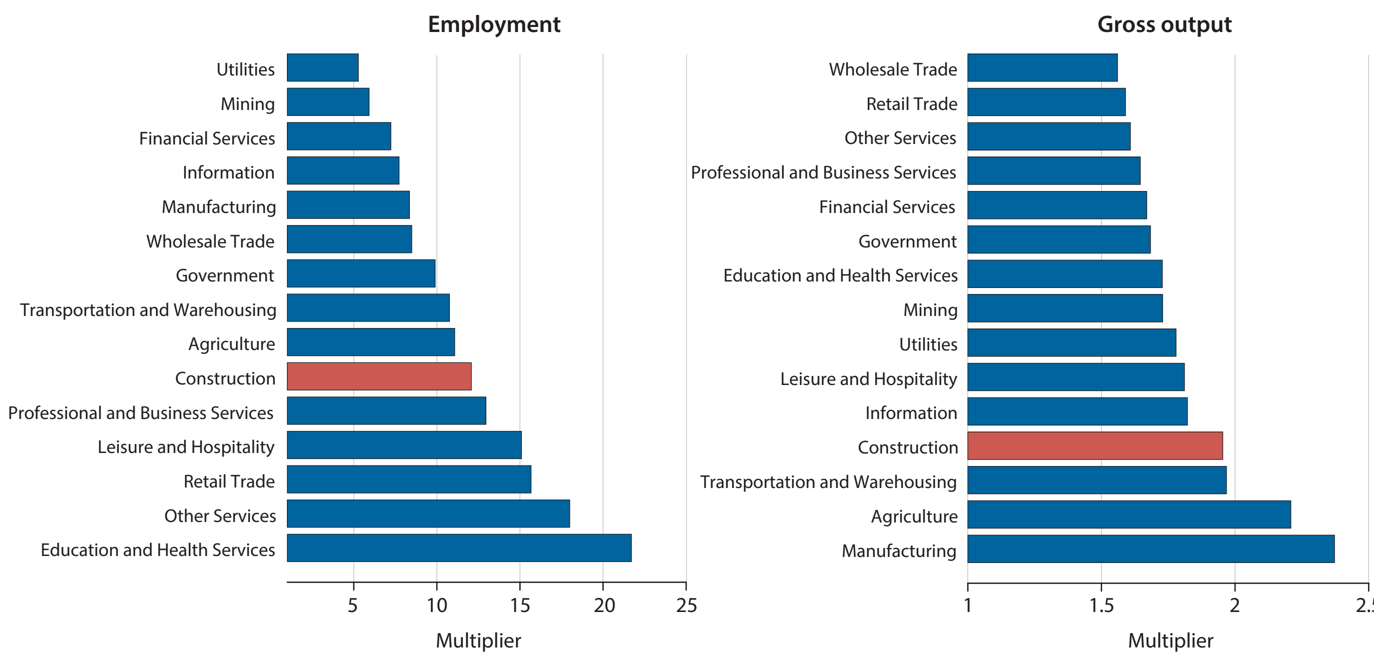
Purchases from Other Sectors



SOURCE: Use matrix from the 2006 BEA input-output tables.

Figure 3

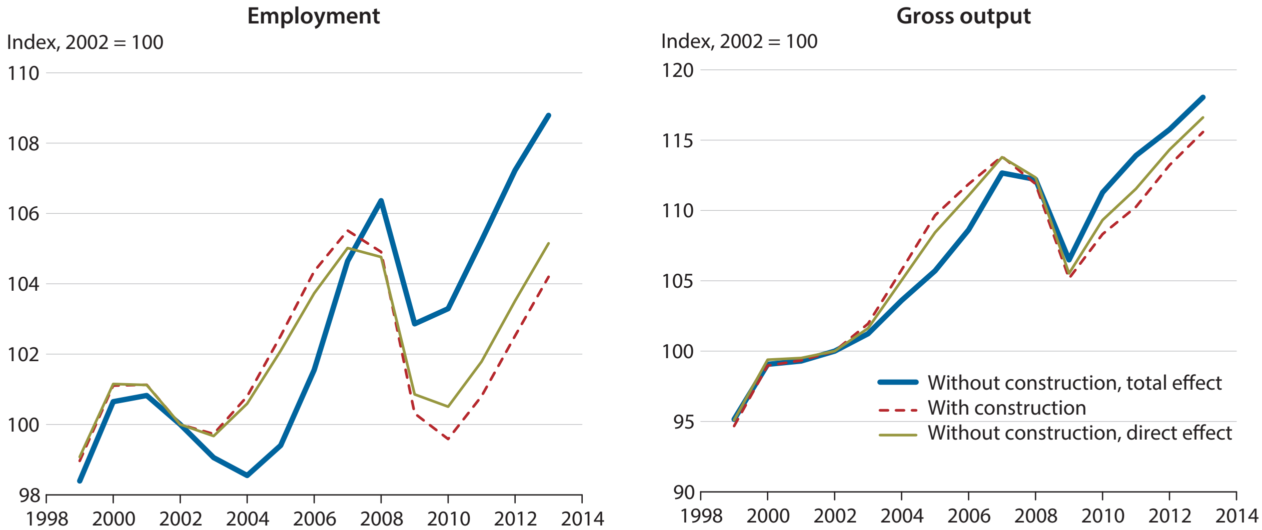
Sectors' Multipliers



SOURCE: Use matrix from the 2006 BEA and BLS input-output tables.

Figure 4

The Construction Sector's Contribution to the Dynamics of Employment and Gross Output



SOURCE: Authors' calculations using BEA and BLS requirements tables.

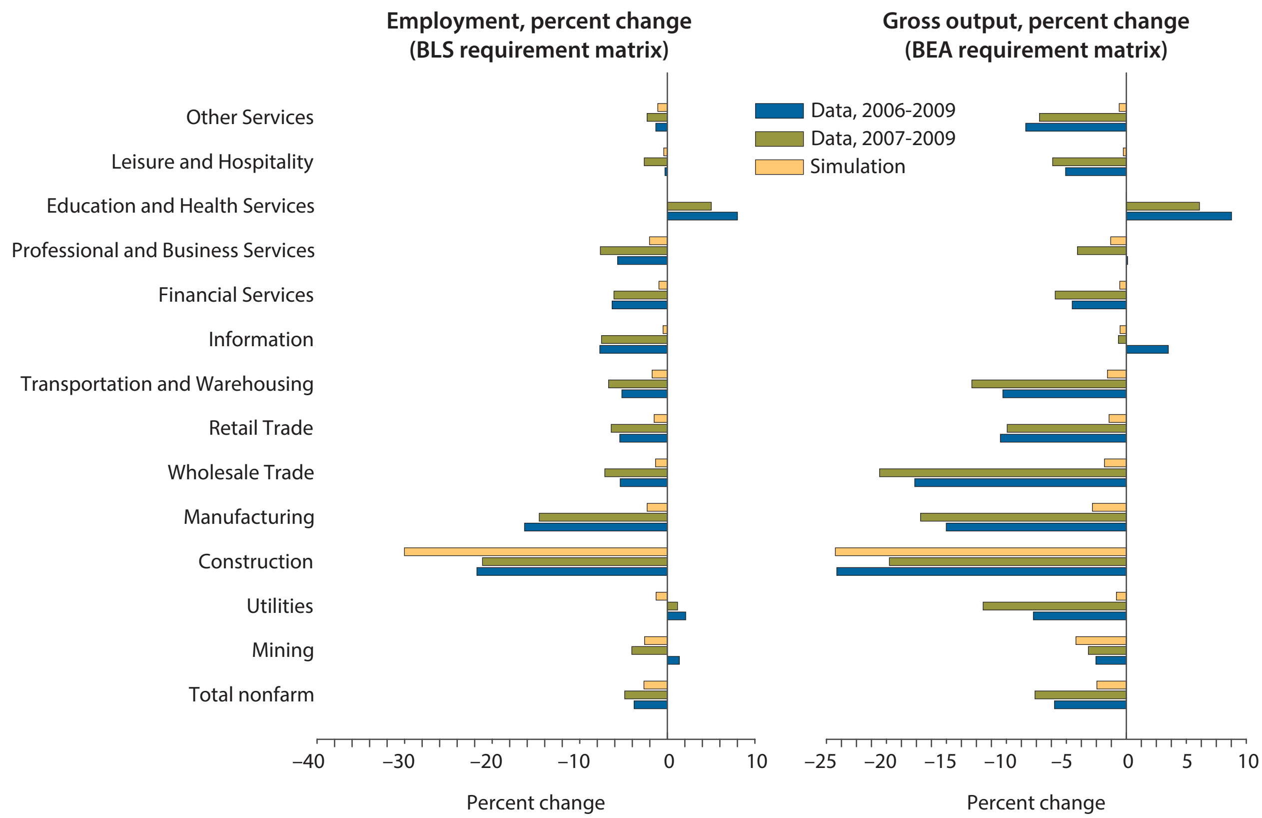
can be used to calculate the *total effect* of a sectoral shock on the rest of the economy. Figure 3 ranks sectors according to the size of their multipliers in terms of gross output and employment.

In terms of gross output, the two sectors with the largest multipliers are manufacturing (2.4) and agriculture (2.3). Construction has the fourth largest production multiplier: A $1 decline in the output of the construction sector generates (absent changes in relative prices and in the composition of final demand) a slightly below $2 decline in gross output of all other sectors combined. Recall that the construction sector is larger than the agriculture sector, its final demand much more volatile than those of both the manufacturing and agriculture sectors, and its output composition more homogeneous. In terms of employment, the construction sector also has a relatively large multiplier. It is worth noting that with respect to employment, the multipliers of the manufacturing and agriculture sectors are not as significant as the multiplier for the construction sector. This highlights that the construction sector is important overall because of its employment and gross output multipliers.

Our goal is to understand how significant these multipliers are when it comes to aggregate fluctuations. To this end, we add up the direct and multiplier effects for the construction sector to compute the *total effect of construction on the rest of the economy*. We use the requirement matrices and compare the actual evolution of U.S. employment and gross output with a counterfactual economy without the construction sector. Figure 4 displays the paths of employment and gross output for the three cases: the actual values and the values without the direct and without the total effects of construction. The difference between these paths is a rough estimate of the construction sector's (direct and total) impact on the aggregate dynamics. When construction is included, total employment increases about 6 percent between 2002 and 2006,

Figure 5

Sectoral Changes During the Great Recession: The Data and Input-Output Simulations



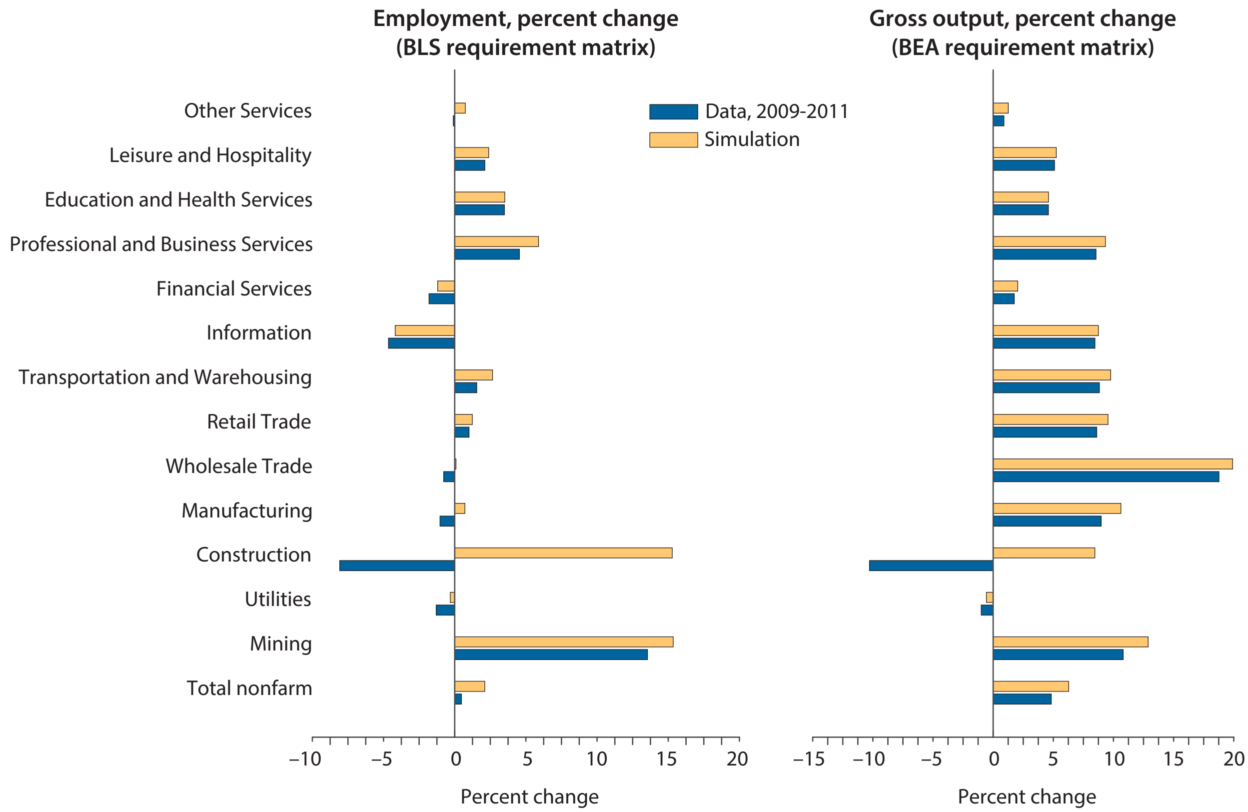
SOURCE: Authors' calculations using BEA data.

which is then entirely lost. In contrast, the economy without a construction sector and with the total effect has a slower recovery from the 2000-01 recession; employment growth picks up only in 2005 and employment destruction starts in 2009. The magnitude of the subsequent decline is half that actually experienced. Unlike in the actual economy, employment starts recovering already in 2010 and in 2012 surpasses the previous peak. This exercise shows that the construction sector contributed greatly to employment growth between 2002 and 2005 and to employment destruction during the Great Recession. A similar, if somewhat weaker, conclusion can be drawn by analyzing the panel for gross output. This simple decomposition reveals that during the Great Recession construction alone may have accounted for 52 percent of the decline in employment and 35 percent of the decline in gross output.

At a more micro level, construction interacts differently with the various sectors in the economy. Therefore, a decline in the activity of the construction sector will have a larger impact on those sectors that sell to it directly as opposed to those that do not. To show this, Figure 5 reports for each sector the actual sectoral declines in employment and gross output, respectively, between 2006-09 and 2007-09 and those estimated using the input-output matrix for

Figure 6

Sectoral Changes During the Recovery: The Data and Input-Output Simulations



SOURCE: Authors' calculations using BEA data.

2007-09, as a consequence of the observed decline in construction.[7] The blue (2006-09) and green bars (2007-09) in Figure 5 represent the historical percent changes in gross output and employment for 13 industries and for the total economy (total nonfarm). In 2006-09, gross output in the construction sector and in the aggregate (total nonfarm) declined close to 25 percent and 6.2 percent, respectively. Employment in the construction sector decreased by roughly the same amount as gross output, 21.5 percent, while aggregate employment declined by 4.4 percent. The aggregate numbers are slightly larger when considering the period 2007-09. The yellow bars represent the declines attributable to the construction sector on the basis of the input-output multipliers (for gross output and employment, respectively). For example, according to this methodology, the drop in the construction sector accounts for a significant part of the gross output decline in the mining sector, about 68 percent of it, while it accounts for little of the decline in the retail trade sector.

According to this methodology, construction is capable of accounting for about 35 percent of the decline in aggregate gross output and for about 52 percent of the decline in aggregate employment. These numbers contrast with the direct impact estimates that account for 20.3

percent of the decline in employment and 19.3 percent of the decline income, as shown in Table 1. The difference between the direct and the total effects of construction is due to the magnifying role of the production interlinkages, and this is what we label the production multiplier.

The construction sector played an important role not only in the Great Recession but also in the subsequent slow recovery. Its contribution can be measured by performing a similar counterfactual for the period 2009-11. Figure 6 shows the simulated growth rates for the 13 sectors and the total economy (nonfarm economy) under the assumption that construction grows from 2009 forward at pre-recession rates. The blue bars display the actual percent changes of gross output and employment, respectively. Between 2009 and 2011, gross output increased by 5 percent and employment increased by roughly 1 percent. The yellow bars represent the counterfactual simulation and show that if construction had grown at its pre-recession levels, total gross output and employment would have increased by 6 percent and 2 percent, respectively. In this scenario, the sectors that would have grown the most in terms of gross output are wholesale trade (20 percent), retail trade (10 percent), mining (13 percent), and transportation and warehousing (11 percent). The differences between the above growth rates and the actual growth rates (blue bars) indicate that the contribution of the construction sector to the dynamics of aggregate employment and output is nontrivial. The next section proposes a simple model of interlinkages that explains the nature of these effects.

4 THEORETICAL MODEL OF INPUT-OUTPUT LINKAGES

This section presents a stylized two-sector model with housing and production interlinkages. The model also incorporates the durable consumption goods nature of housing, the presence of a fixed factor in the production of housing services (i.e., land), and the complementarity between consumption of housing services and of all other goods.

4.1 Households

Total population size, N_t, is normalized to 1. Household preferences are defined by a time-separable utility function, $U(c_t, \theta_t h_t, n_t)$, where c_t represents consumption goods, h_t represents housing services, and n_t represents labor supplied in the market. Housing provides utility, and it is a complement to aggregate consumption. The shifts in housing consumption are driven by adjustments in the parameter θ_t. The utility function U satisfies the usual properties of differentiability and concavity. The sequence of utilities is discounted by the term $\beta \in (0,1)$. Housing services are obtained by combining physical structures, s_t, and land, l_t, according to $H(s_t, l_t)$. The latter is homogeneous of degree 1 and satisfies $H_i' > 0$, $H_i'' > 0$ and $H_{ij}'' > 0$. Housing structures depreciate at a constant rate, δ_s. In each period, the numeraire is the spot price of the nonconstruction good. Formally, the representative consumer chooses $\left\{c_t, h_t, n_t, k_{t+1}, s_{t+1}, l_{t+1}\right\}_{t=0}^{\infty}$ to maximize

$$\max \sum_{t=0}^{\infty} \beta^t U\left(c_t, \theta_t h_t, n_t\right),$$

$$s.t. \qquad c_t + x_t^k + p_t^s x_t^s = w_t n_t + r_t^k k_t + p_t^l \left(l_t - l_{t+1}\right) + \pi_t,$$

(1) $$h_t = H(s_t, l_t),$$

(2) $$x_t^k = k_{t+1} - (1-\delta_k)k_t \geq 0,$$

(3) $$x_t^s = s_{t+1} - (1-\delta_s)s_t \geq 0.$$

The maximization is also subject to transversality and no-Ponzi-game conditions. Prices are defined as follows: p_t^s is the price of infrastructure, p_t^l is the price of land, w_t is the wage rate, and r_t is the gross return on capital. To facilitate computing the rental rate for housing services, our specification allows land trading, l_t, even if in equilibrium there is no trading of land, which is owned by the representative household and inelastically supplied. The term π_t represents profits from the construction sector. All investment decisions are subject to an irreversibility constraint and have different depreciation rates in the two sectors, construction and nonconstruction.

The relevant first-order conditions of the consumer problem are

$$\frac{U_n(c_t,\theta_t h_t,n_t)}{U_c(c_t,\theta_t h_t,n_t)} = w_t, \qquad \forall t,$$

$$\frac{U_c(c_t,\theta_t h_t,n_t)}{\beta U_c(c_{t+1},\theta_{t+1}h_{t+1},n_{t+1})} = 1 + r_{t+1}^k - \delta_k, \qquad \forall t,$$

when the irreversibility constraints do not bind, $x_t^k > 0$. The relevant conditions for housing decisions for the case with positive housing investment ($x_t^s > 0$) satisfy

$$\frac{U_h(c_t,\theta_t h_t,n_t)}{U_c(c_t,\theta_t h_t,n_t)} = R_t, \qquad \forall t,$$

$$p_t^s = \frac{1}{1+r_{t+1}^k}\left[R_{t+1}H_s(s_t,l_t) + p_{t+1}^s(1-\delta_s)\right],$$

$$p_t^l = \frac{1}{1+r_{t+1}}\left[R_{t+1}H_l(s_t,l_t) + p_{t+1}^l\right],$$

where R_t represents the implicit rental price for housing services measured in terms of consumption units. Notice that a no-arbitrage condition holds between investment in land and housing. The last two expressions state that the current cost of purchasing a unit of housing structures (land) equals the future return of housing services derived from the housing capital (land) valued at market prices, plus its capitalization.

4.2 Nonconstruction Sector

The model uses a 2×2 input-output structure: To operate, each sector requires, among other things, that the output of a sector uses intermediate inputs from other sectors as well as its own. To capture this fact, we deviate from common practice and write all production functions in terms of gross (as opposed to net, i.e., value-added) output. Capital goods, which are produced in the nonconstruction sector, must be distinguished from the intermediate inputs from the same sector since they last more than one period. In the baseline model, capital goods are used only in the nonconstruction sector for simplicity. Both investments satisfy the putty-clay assumption, which is sector specific.

Formally, let $m^{i,j}$ be the intermediate input produced by sector i and used by sector j. The nonconstruction sector operates in a competitive market and uses the technology $A_t^y F\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right)$ to produce its gross output:

$$Y_t = c_t + x_t^k + m_t^{y,y} + m_t^{y,s}.$$

The production function F has constant returns to scale. The firm's optimization problem is

$$\pi_t^y = \max_{k_t, n_t^y, m_t^{y,y}, m_t^{s,y}} Y_t - w_t n_t^y - r_t^k k_t - m_t^{y,y} - p_t^s m_t^{s,y}, \qquad \forall t,$$

$$s.t. \qquad Y_t = A_t^y F\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \qquad \forall t,$$

where the price of the nonconstruction sector's output is normalized to 1. The constant-returns-to-scale assumption implies zero equilibrium profits, $\pi_t^y = 0$, and marginal cost pricing for each input

$$\begin{aligned}
r_t^k &= A_t^y F_1\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \\
w_t &= A_t^y F_2\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \\
1 &= A_t^y F_3\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \\
p_t^s &= A_t^y F_4\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right).
\end{aligned}$$

4.3 Construction Sector

The construction sector is also competitive. Its net output consists of residential structures, purchased by the households, while its gross output also includes structures used as intermediate inputs in both sectors. In the baseline case, purely for simplicity, we assume this sector has a fixed stock of capital; hence its value added is split between the wages of labor and the rent accruing to the owner of the fixed capital stock (the representative household). Implicit in this formulation is a somewhat extreme assumption about the mobility of factors from one sector to another: While labor can move freely, the stock of capital invested in the construction sector is completely immobile (either way), and variations in investment activity have an impact only on the nonconstruction sector. The technology for gross output is represented by

$$X_t^S = x_t^s + m_t^{s,s} + m_t^{s,y} = A_t^s G\left(n_t^s, m_t^s\left(m_t^{s,s}, m_t^{y,s}\right)\right)$$

and exhibits decreasing returns to scale in labor and the intermediate input mix. The function $G(\cdot)$ has a constant elasticity of substitution, and the aggregator of intermediate inputs is homogeneous of degree 1. The optimization problem of the representative firm is now

$$\pi_t^s = \max_{n_t^s, m_t^{s,s}, m_t^{y,s}} p_t^s X_t^S - w_t n_t^s - p_t^s m_t^{s,s} - p_t^y m_t^{y,s}, \qquad \forall t,$$

$$s.t. \qquad X_t^S = A_t^s G\left(n_t^s, m_t^s\left(m_t^{s,s}, m_t^{y,s}\right)\right), \qquad \forall t.$$

The first-order conditions are similar to those of the representative firm in the nonconstruction sector and are not repeated here. Note that because of the presence of a fixed stock of capital, firm profits are not zero in equilibrium in this sector. It is worth emphasizing that p_t^s reflects the cost of producing new structures. The equilibrium price of a house differs from this value since it depends on the relative value of the structures and land.

4.4 Competitive Equilibrium

The competitive equilibrium of this economy is defined as follows:

Competitive Equilibrium: *Given a sequence of values* $\left\{A_t^y, A_t^s, \theta_t\right\}_{t=0}^{\infty}$, *a competitive equilibrium consists of allocations* $\left\{c_t, x_t^k, x_t^s, l_t, n_t^y, n_t^s, m_t^{s,s}, m_t^{y,s}, m_t^{y,y}, m_t^{s,y}\right\}_{t=0}^{\infty}$ *and prices* $\left\{w_t, r_t^k, p_t^l, p_t^s, r_t, R_t\right\}_{t=0}^{\infty}$ *that satisfy the following:*

(i) the household's optimization problem;

(ii) profit maximization in the construction and nonconstruction sectors; and

(iii) the clearing of markets:

(a) the labor market (w_t)*:*

$$n_t = n_t^y + n_t^s, \qquad \forall t;$$

(b) the land market (p_t^l)*:*

$$l_t = l_{t-1} = \overline{l}, \qquad \forall t;$$

(c) the capital market (r_t^k)*:*

$$r_t^k = A_t^y F_1\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \qquad \forall t;$$

(d) the nonconstruction output market ($p_t^c = 1$)*:*

$$c_t + x_t^k + m_t^{y,y} + m_t^{y,s} = A_t^y F\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \qquad \forall t; \text{ and}$$

(e) the construction output market (p_t^s)*:*

$$x_t^s + m_t^{s,s} + m_t^{s,y} = A_t^s G\left(n_t^s, m_t^s\left(m_t^{s,s}, m_t^{y,s}\right)\right), \qquad \forall t.$$

For a given sequence of housing demand shifters $\{\theta_t\}$ the model endogenously generates time series for all macroeconomic quantities and prices. Before comparing the predictions of the model with the data, it is useful to understand how to characterize the amplification process. This is described in the next subsection.

4.5 The Macroeconomic Effects of Production Interlinkages and Demand Complementarities

To proceed analytically it is necessary to make a number of simplifying assumptions relative to the baseline model: (i) the economy lasts one period, (ii) labor and intermediate goods are the only inputs, and (iii) the share of land in the housing aggregator is zero. The utility index is defined as $U(c, \theta, h, n)$, and the budget constraint of the representative household is $c + ph = wn$, where w is the wage rate and p the price of housing, both measured in units of the nonconstruction good.

Part of the output of the nonconstruction sector, m_y, is used as an input to produce construction, and part of the output of the construction sector is used to produce nonconstruction goods, m_h. The gross output flows of the two sectors are $c + m_y = Y = A_y f(n_y, \varepsilon_y m_h)$ and $h + m_h = H = A_h g(n_h, \varepsilon_h m_y)$, respectively, where A_j represents the productivity of sector $j = y,h$.[8] The $\varepsilon_j (j = y,h)$ terms capture the relative importance in sector j of the intermediate inputs from the other sector. Aggregate labor satisfies the restriction $n_y + n_h = n$. Free mobility implies that the wage rate is the same across sectors.

A *competitive equilibrium* in this economy is an allocation $\{c, h, n_y, n_h, m_y, m_h\}$ and prices $\{w, p\}$ that solve (i) the optimization problem of the household, (ii) the optimization problem of the firms in each sector, and (iii) the market-clearing conditions.

As a function of the preferences (θ) and technology (ε) parameters, value added in this economy is defined as

$$VA(\theta,\varepsilon) = c(\theta,\varepsilon) + p(\theta,\varepsilon)h(\theta,\varepsilon).$$

The goal is to identify conditions under which a shift $\Delta\theta = \theta' - \theta$ in the demand for housing has a larger impact on total employment and value added when there are interlinkages ($\varepsilon > 0$) rather than when there are not ($\varepsilon = 0$), that is, the conditions under which we have

$$\frac{\partial VA(\theta,\varepsilon)}{\partial\theta} \geq \frac{\partial VA(\theta,0)}{\partial\theta}.$$

There are three interacting channels through which a change in the demand for housing may affect value added: (i) a direct change in the desired quantities of c and h, (ii) a change in their relative prices (and the consequent second-order changes in the quantities demanded), and (iii) a change in the supply of labor due to wealth and price effects. To highlight the difference of this transmission mechanism relative to the recent literature discussed in Section 2 above, the next examples abstract from movements in relative prices due to sectoral shocks. The full quantitative model will also consider changes in relative prices and the related adjustments in quantities.

4.5.1 Example: Leontief Production. A simple way to eliminate the price effects is to consider an economy in which both production functions have fixed coefficients

$$c+m_y=Y=A_y\min\left\{n_y,\frac{m_h}{\varepsilon_y}\right\},$$

$$h+m_h=H=A_h\min\left\{n_h,\frac{m_y}{\varepsilon_h}\right\}.$$

The parameters $\varepsilon_y \geq 0$ and $\varepsilon_h \geq 0$ capture the intensity of the sectoral interlinkages.[9] Using the nonsubstitutability of inputs and the constraint on total employment yields a linear production possibility frontier:

$$c+\left(\frac{A_y+\varepsilon_h}{A_h+\varepsilon_y}\right)h=\left(\frac{A_yA_h-\varepsilon_y\varepsilon_h}{A_y+\varepsilon_h}\right)n.$$

To satisfy feasibility, it must be the case that $A_yA_h > \varepsilon_y\varepsilon_h$. If the intermediate input requirements are too high relative to the productivity of each sector, it would not be feasible to produce positive amounts of both goods.[10] The linearity of the production possibility frontier implies that the relative price of new houses depends only on the technical coefficients,

$$p=\frac{A_y+\varepsilon_h}{A_h+\varepsilon_y}.$$

In the model without interlinkages $\varepsilon_j \to 0$, the price is given by the ratio of productivities $p = A_y/A_h$. Similarly, wages are determined by

$$w=\left(\frac{A_yA_h-\varepsilon_y\varepsilon_h}{A_y+\varepsilon_h}\right).$$

In the absence of sectoral shocks on (A_y,A_h), exogenous changes in housing demand $(\Delta\theta)$ have no effect on prices and wages. All the macroeconomic effects of value added are driven by changes in the production of each sector:

$$VA(\theta,\varepsilon)=c(\theta,\varepsilon)+\left(\frac{A_y+\varepsilon_h}{A_h+\varepsilon_y}\right)\cdot h(\theta,\varepsilon).$$

From this expression, it is direct to derive specifications for which a shift in housing demand is *not amplified* in the aggregate. The first one ignores sectoral interlinkages, $\varepsilon_y = \varepsilon_h = 0$ as the model collapses to the standard two-sector model where relative prices are determined by factor productivities. The second one assumes perfectly symmetric sectors, $\varepsilon_y = \varepsilon_h = \varepsilon$ and $A_y = A_h = A$, implying steady-state prices and wages equal to $p = 1$ and $w = (A - \varepsilon)$. One can easily add a third trivial case where the nonconstruction good is completely independent of θ; that is, $c(\theta,\varepsilon) = c(\varepsilon)$.

Case 1: Preferences with Perfect Complementarity. This specification allows the housing-demand shifter $\Delta\theta$ to change directly the consumption demand of both goods. The utility index is given by $U(\theta c, h, n) = \min\{\theta c, h\} - an^{1+\gamma}/(1 + \gamma)$. This corresponds to the extreme case of perfect complementarity; but only some degree of complementarity (less than unitary elasticity of substitution between c and h) is sufficient for the mechanism to operate. With this utility function consumption is given by

$$\hat{c} = \left[\frac{A_y A_h - \varepsilon_y \varepsilon_h}{\left(A_h + \varepsilon_y\right) + \theta\left(A_y + \varepsilon_h\right)} \right] n,$$

and the demand for housing is just $\hat{h} + \theta\hat{c}$. In the model without interlinkages ($\varepsilon_j = 0$), employment is allocated in the two sectors according to $\hat{n}_h^{nolink} = \left(A_y / \left(A_y + \theta A_h\right)\right) n$.

Solving for the aggregate level of employment yields

$$\hat{n}^{link} = \left[\frac{\theta}{a} \left(\frac{A_y A_h - \varepsilon_y \varepsilon_h}{\left(A_h + \varepsilon_y\right) + \theta\left(A_y + \varepsilon_h\right)} \right) \right]^{\frac{1}{\gamma}},$$

whereas in the absence of interlinkages the employment level is

$$\hat{n}^{nolink} = \left[\frac{\theta}{a} \left(\frac{A_y A_h}{A_h + \theta A_y} \right) \right]^{\frac{1}{\gamma}}.$$

Measured economic activity is given by

$$VA = c + ph = \left[\frac{A_y A_h - \varepsilon_y \varepsilon_h}{A_y + \varepsilon_h} \right] \hat{n}^i(\theta).$$

Notice that value added is proportional to total employment, and the scaling factor does not depend on the parameter θ. The change in value added due to a change in housing demand driven by $\Delta\theta$ is

$$\frac{\partial VA}{\partial \theta} = \left[\frac{A_y A_h - \varepsilon_y \varepsilon_h}{A_y + \varepsilon_h} \right] \frac{\partial \hat{n}(\theta)}{\partial \theta}.$$

We ask next, how do changes in the preference parameter θ affect aggregate employment, n, and value added, VA? Notice first, from the formulas above, that the economy with interlinkages and the one without them have different levels of aggregate employment. Hence, we will compute the two elasticities of employment with respect to variations in θ.

In the economy with linkages, this elasticity is

$$\epsilon_{n,\theta}^{link} = \frac{1}{\gamma} \left[\frac{A_y + \varepsilon_h}{\left(A_h + \varepsilon_y\right) + \theta\left(A_y + \varepsilon_h\right)} \right] > 0,$$

and in the economy without linkages it is

$$\epsilon_{n,\theta}^{nolink} = \frac{1}{\gamma} \left[\frac{A_y}{A_h + \theta A_y} \right] > 0.$$

The presence of interlinkages amplifies the effect of any given preference shock when $\epsilon_{n,\theta}^{link} > \epsilon_{n,\theta}^{nolink}$, which reduces to $A_h \varepsilon_h > A_y \varepsilon_y$ after a bit of algebra.[11] This condition is clearly satisfied when the construction sector purchases intermediate inputs from the rest of the economy ($\varepsilon_h > 0$), but not the other way around ($\varepsilon_y = 0$). Hence, in general, the condition holds (for given levels of sectoral productivities) when the construction sector absorbs lots of inputs from the other sector but the other sector does not use housing as an intermediate input, which does not sound so unrealistic. Notice that when there is symmetry between both sectors the condition fails. This is consistent with the earlier theoretical results of Horvath (1998) and Dupor (1999) cited in Section 2.

Is there some empirical evidence that supports this asymmetry? We have used the direct-input requirement matrices to carry out a back-of-the-envelope test of these conditions. Our procedure was simple: We aggregated the matrices into a 2×2 format: construction and everything else. Next, we eliminated the "own intermediate inputs," which the model assumes away for simplicity, and collapsed all the value added of the two sectors into labor income. The sectoral price indices were used to compute the relative price p during the available sample period and then to compute, by simple algebra, the four parameters of our model. We found that $A_h \varepsilon_h / A_y \varepsilon_y$ equals 636, after rounding up. The inequality is amply satisfied, thereby suggesting, on the basis of this admittedly simplified model, that in the real world the magnification effects of asymmetries are likely to be present.

Case 2: Preferences with Unitary Elasticity. To highlight the importance of the complementarity between c and h, we consider the case of Cobb-Douglas preferences, $u(c,h) = \log c + \theta \log h$. This specification is very common in macro housing models (i.e., Davis and Heathcote, 2005, and Iacoviello, 2005) and in the recent literature on production networks (i.e., Acemoglu et al., 2012). It implies a total employment level

$$\hat{n} = \omega^i \left((1+\theta)/a\right)^{\frac{1}{1+\gamma}},$$

where $\omega^{nolink} = 1$ in the economy without linkages and $\omega^{link} > 1$ in the economy with them, independently of θ. The level of aggregate employment is larger in the economy with linkages than in the one without them; but in response to changes in housing ($\Delta\theta$), the two economies share the same aggregate labor elasticity,

$$\epsilon_{n,\theta}^{link} = \epsilon_{n,\theta}^{nolink} = \frac{1}{1+\gamma}\left(\frac{\theta}{1+\theta}\right).$$

Even with Leontief technologies, the unitary elasticity eliminates the contribution of the input-output multipliers. For this reason, in the quantitative exercise in Section 5, we explore the importance of demand complementarities by adopting a more general class of preferences that includes the Cobb-Douglas as a special case.

5 QUANTITATIVE ANALYSIS

In this section we use the general model developed in the previous section to carry out various quantitative exercises.

5.1 Parameterization

To proceed we need to specify functional forms and then assign parameter values. The choice of functional forms is relatively general with the exception of the housing demand shifter,

$$u(c_t,h_t,N_t)=\frac{\left[\eta c_t^{-\rho}+(1-\eta)h_t^{-\rho}\right]^{-\frac{1-\sigma}{\rho}}+\theta_t h_t}{1-\sigma}+\varrho\log(1-N_t),$$

where the parameter ρ controls the degree of demand complementarity between consumption, c, and housing services, h. The parameter σ represents the intertemporal elasticity of substitution, and the parameter η represents the relative importance of consumption. The utility from leisure is logarithmic, as is standard in the real business cycle literature with a representative agent.[12] Housing services enter as complementary to consumption goods but also as a linear term scaled by the parameter θ_t that shifts housing demand to generate a construction boom and bust. This is interpreted as a "reduced-form housing market wedge" given by $\tilde{\theta}_t$ that, when measured relative to consumption goods, implies the housing price equation

$$p_t^h=R_t+\tilde{\theta}_t+\frac{p_{t+1}^h}{1+r_{t+1}},$$

where the notion of rents comes from owner-equivalent rent given by

$$R_t=\frac{(1-\eta)}{\eta}\left(\frac{c_t}{h_t}\right)^{1+\rho}.$$

Housing services are obtained from housing structures and land according to a Cobb-Douglas mixture,

$$h=H(s,l)=z_h(s)^{\epsilon}\left(\bar{l}\right)^{1-\epsilon},$$

where z_h represents a transformation factor between stock and flow. The production of the nonconstruction goods also uses a Cobb-Douglas technology:

$$F\left(k,n^y,m^{y,y},m^{s,y}\right)=A^y(k)^{\alpha_1}\left(n^y\right)^{\alpha_2}\left(m^{y,y}\right)^{\alpha_3}\left(m^{s,y}\right)^{1-\alpha_1-\alpha_2-\alpha_3},$$

where α_i represents the share in production for input i. Notice that the specification allows for substitutability between intermediate goods. The technology used in the construction sector, instead, is a constant elasticity of substitution with diminishing returns to scale:

$$G\left(n^s,m^{s,s},m^{y,s}\right)=A^s\left[\gamma_2\left(n^s\right)^{-\gamma_1\gamma_4}+(1-\gamma_2)\left(\left(m^{s,s}\right)^{\gamma_3}\left(m^{y,s}\right)^{1-\gamma_3}\right)^{-\gamma_4}\right]^{-1/\gamma_4}.$$

Table 2

Parameter Values

Parameter	α_1	α_2	α_3	γ_1	γ_2	γ_3	γ_4	A^y	A^s	z_h	ε	η	σ	ρ
Value	0.18	0.50	0.035	0.62	0.40	0.04	1.5	2.4	1.74	0.175	0.28	0.435	1	5

The model parameters are set to match long-run averages of their data counterparts between 1952 and 2000. The implied parameter values are relatively robust to the choice of the sample period; however, during the housing boom some of the ratios and long-run averages departed significantly from their historical trends. Hence, to avoid stacking the cards in favor of our model, we used data from only the period before the housing boom and bust to calibrate our model.

The time unit is a year, as input-output tables are yearly at best. The discount factor is $\beta = 0.96$. The depreciation rates of residential structures and nonresidential capital are $\delta^s = 0.015$ and $\delta^y = 0.115$, respectively. The weight on leisure, $\varrho = 0.33$, is such that total hours worked equal one-third of the time endowment in the steady state. The preference parameters are set to match the consumption-to-output and housing-to-output ratios. The parameters of the production functions are set to satisfy the following:

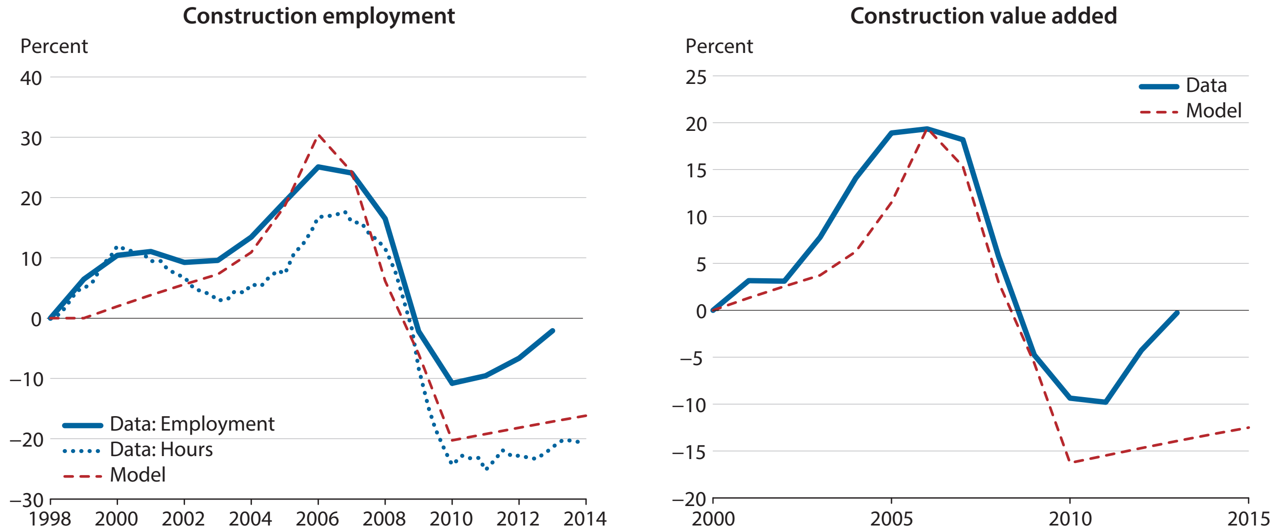
(i) the ratio of gross output in the two sectors, $Y^s/Y^y = 0.08$;
(ii) the average labor share in the construction sector = 0.7;
(iii) the average labor share in the nonconstruction sector = 0.65;
(iv) the ratio of consumption to nonconstruction gross output = 0.35;
(v) the observed shares of intermediates in gross output of own sector (m^{yy} and m^{ss}) = 0.4, 0.007;
(vi) time allocated to market activities, $n^y + n^s = 1/3$; and
(vii) the ratio of employment in the two sectors, $n^y/n^s = 16$.

The values of the parameters not listed here are displayed in Table 2. The intratemporal elasticity of substitution between consumption and housing services is determined by the parameter $\varepsilon_{ch} = 1/(1 + \rho)$. Quantitatively, the value of ρ is an important determinant of the spillover effects from housing into the rest of the economy. If consumption services are close substitutes, a decline in the demand for housing services can generate an increase in the demand for the consumption good; whereas if they are close complements, a decline in housing demand can translate into a decline also in the demand for consumption. Various recent papers, part of an extensive literature on the topic, estimate this elasticity to be less than 1. For example, Flavin and Nakagawa (2008) use a model of housing demand and estimate an elasticity of less than 0.2. Others (i.e., Song, 2010, and Landvoigt, 2017) use alternative model specifications and also estimate values for the elasticity to be less than 1. The simulations consider elasticities $\varepsilon_{ch} \in \{0.17, 0.25\}$.

To generate dynamics in housing demand and construction, households face unanticipated shocks to the demand shifter θ_t. Households have some initial expectations about their

Figure 7

Dynamics of the Construction Sector: The Model and Data



SOURCE: BEA and authors' calculations.

housing demand set by the initial values $\theta_t = \theta_{2000}$ for all $t \geq 2000$. Looking forward, they assume this parameter will remain unchanged in the future. In 2001, households are surprised by an initial increase in θ perceived as permanent going forward, $\theta_t = \theta_{2001} > \theta_{2000}$ for all $t \geq 2001$. In each subsequent period, this parameter is adjusted. In 2007, there is a demand reversal that generates a decline in housing demand until 2010; thereafter it remains constant forever.[13] The dynamics of this parameter are calculated to generate equilibrium paths of valued added and employment and hours in the construction sector that are in line with the data, as shown in Figure 7.

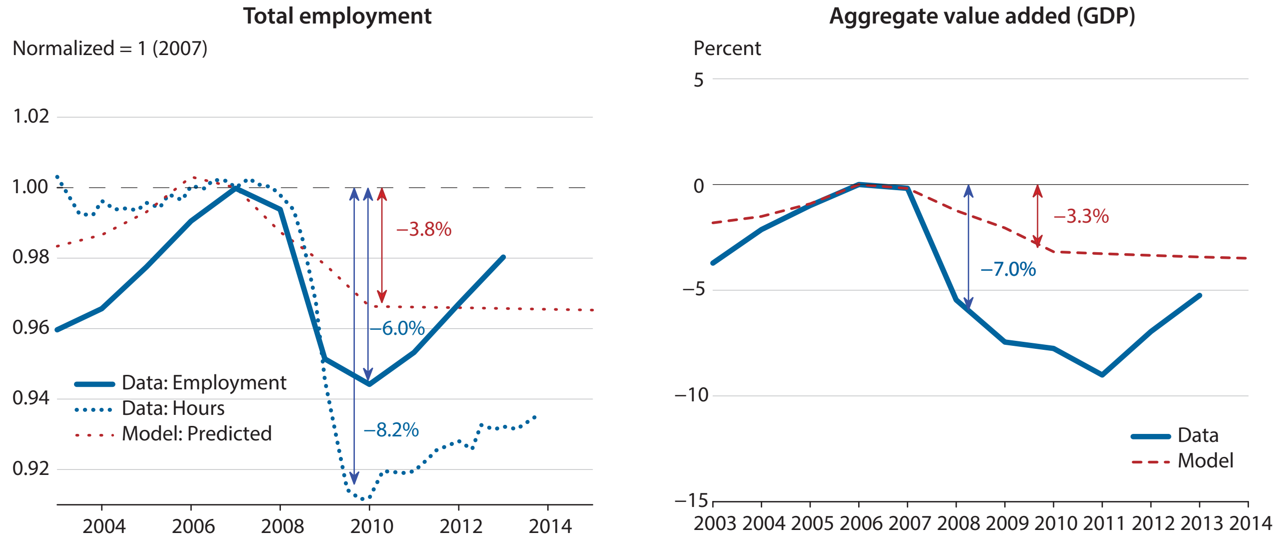
5.2 Role of Residential Investment in Growth and Employment

The goal of this exercise is to measure the sectoral contribution of the construction sector to the macroeconomy. The baseline case considers a boom and a bust in the construction sector, which generates the total employment and aggregate value-added series summarized in Figure 8. The shocks to the construction sector have nontrivial effects on total employment and aggregate value added. To measure the ability of the model to mimic the data, we measure the fraction of the changes in employment and GDP during the expansion period (2000-07) and during the recession (2007-10) it generates. During the boom, the exogenous changes in the demand for housing explain 60 percent of the change in total employment and 25 percent of value added.[14]

Similarly, the housing crash, started by a sudden decrease in housing demand, generates declines in the employment of the construction sector and the demand for intermediate inputs

Figure 8

The Aggregate Implications of a Shock to Construction



SOURCE: BEA and authors' calculations.

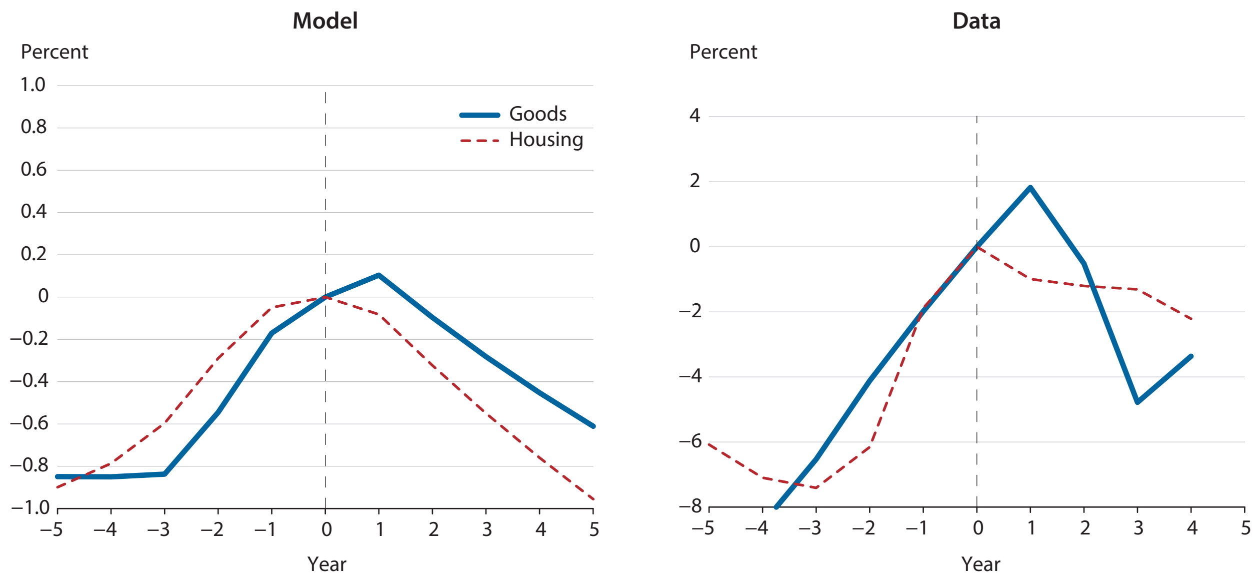
from suppliers. The input-output structure of the model once again lowers the demand for the nonconstruction sector as demand for intermediates from the construction sector falls. In the short run, the decline in the demand for housing generates a very small and short-lived increase in nonhousing consumption that is consistent with the empirical evidence, as shown in Figure 9. This temporary consumption increase is not sufficient to compensate for the decline in other key macroeconomic aggregates. In the model, the collapse of the construction sector (starting in 2007) generates a 3.8 percent decline in total employment and a 3.3 percent decline in aggregate value added. Comparing the numbers with the data, the model can rationalize 44 percent of the decline in employment and 56 percent of the decline in total value added.

An important feature of construction in the business cycle is its leading role during both booms and busts (see Leamer, 2007). Similarly, the data suggest that, during a boom, purchases of housing and durable goods increase faster than purchases of food and services; during a bust, purchases of housing and durable goods decline very sharply, while nonhousing related purchases continue to increase for a few more quarters. Our model captures these lead-lag patterns almost perfectly, as illustrated by the two panels of Figure 9.

In terms of prices, the model generates an 11 percent increase in the house price-to-rent ratio (price-rent ratio) during the housing boom and a 15 percent decline during the bust, as shown in the left panel of Figure 10. According to the OECD data, the price-rent ratio increased 40 percent during the boom and declined 25 percent during the bust. Relative to the data, the model captures around 30 percent of the boom and 60 percent of the bust. The model performs remarkably well given that this is a nontargeted moment and that there is only one force driving

Figure 9

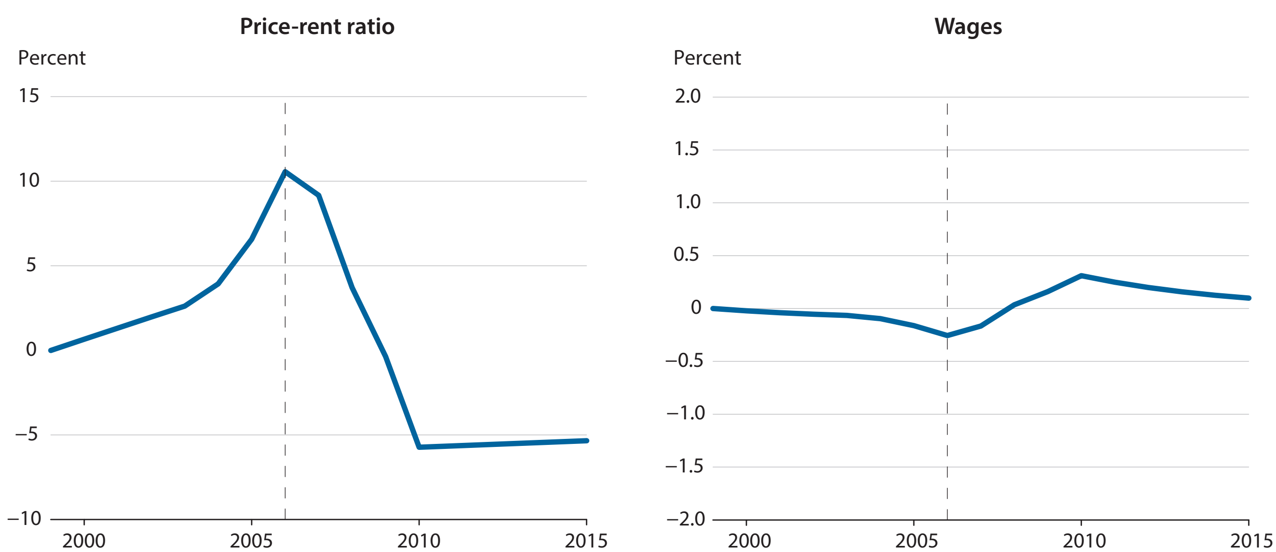
Lead-Lag Responses of Consumption and Housing Spending



SOURCE: BEA and authors' calculations.

Figure 10

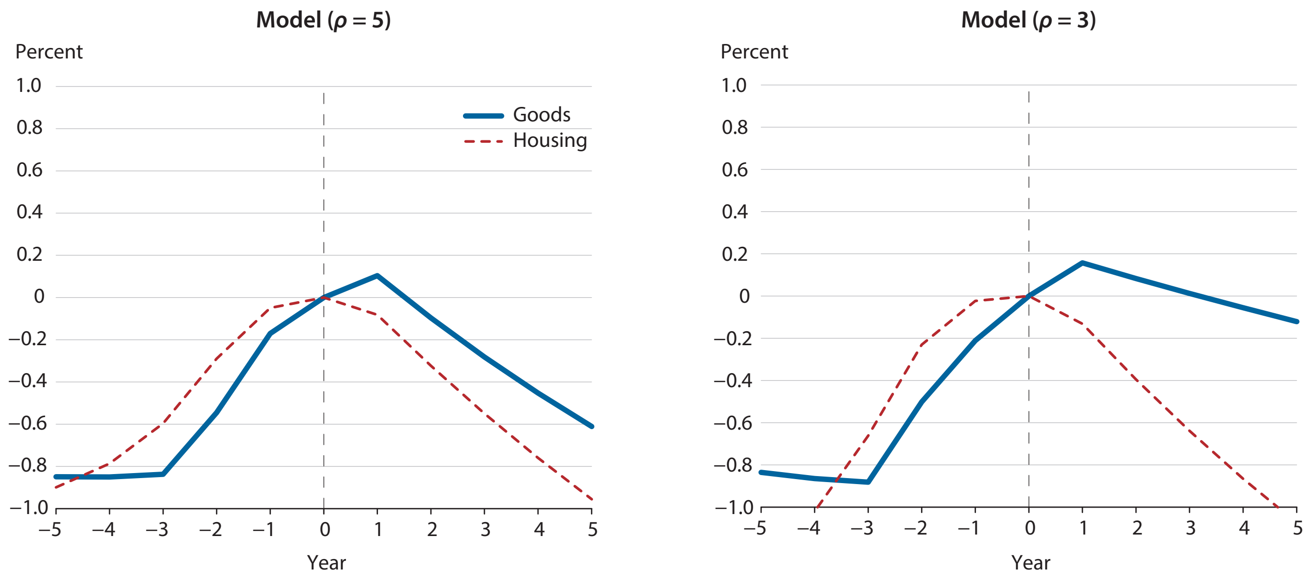
Model-Implied Prices



SOURCE: Authors' calculations.

Figure 11

Lead-Lag Responses of Consumption and Housing Spending



SOURCE: Authors' calculations.

the aggregate dynamics. For a detailed discussion on other drivers for house prices, see Garriga, Manuelli, and Peralta-Alva (2019).

The model can also reconcile large movements in employment and hours with very modest movements in real wages, as shown in the right panel of Figure 10. There is abundant research arguing that a variety of different frictions affecting the labor market are important to generate significant movements in aggregate employment (see, for example, Boldrin and Horvath, 1995; Christiano, Motto, and Rostagno, 2010; Arellano, Bai, and Kehoe, 2019; Gertler and Karadi, 2011; Hall, 2011; and Jones, Midrigan, and Philippon, 2018). Our model abstracts from such features, but modeling the use of intermediates allows us to reduce the response of wages to changes in the labor supply. Section 6 explores the connection of these findings with the existing literature.

5.3 The Role of Demand Complementarities

In our model, the complementarity between consumption and housing is an important driver of the employment dynamics. To compare the model implications for different levels of complementarity, we calculate, in each case, a sequence of demand shifters $\{\theta_t\}$ that match the dynamics of employment in the construction sector. The qualitative implications are the same, but the calibration with stronger complementarity generates a more pronounced boom-bust pattern. With a lower degree of complementarity, GDP falls 2 percent instead of 3.3 percent and total employment declines by 2.4 percent instead of 3.8 percent. The degree of complementarity also has implications for other variables. Figure 11 emphasizes the different lead-lag responses of consumption and housing spending. When the elasticity of substitution between consumption and housing is increased from 0.16 to 0.25, consumption declines significantly less during a bust. With low elasticity, expenditures on goods respond with a lag relative to housing but have similar dynamics. As the elasticity increases, the dynamics of goods expenditures diverge from the declining path of housing. Increasing the elasticity to higher numbers ($\rho < 2$) would generate a boom of the nonconstruction sector during the collapse of the construction sector. A low elasticity of substitution magnifies the aggregate responses of employment and capital investment.

5.4 The Role of Interlinkages and Supply Complementarities

In our view the interlinkages have been an important driver of aggregate output and employment during the housing boom and bust. To isolate the effects of the interlinkages from those derived purely from consumer demand for housing, we study two alternative specifications. The first uses the same parametric calibration and shuts off interlinkages by holding the sectoral demand of intermediates fixed at the level of the initial steady state in 1998 $\left(m_t^{s,s} = m_0^{s,s}, m_t^{y,s} = m_0^{y,s}, m_t^{y,y} = m_0^{y,y}, m_t^{s,y} = m_0^{s,y}\right)$. This case is referred to as the "no interlinkages specification." The second formulation completely ignores the role of intermediate goods $\left(m_t^{s,s} = m_t^{y,s} = m_t^{y,y} = m_t^{s,y} = 0\right)$, and the production functions are specified for value added and not gross output.[15] In this "value-added specification," the relevant technologies ignore the use of intermediate inputs from other sectors in the economy. The nonconstruction sector satisfies

$$c_t + x_t^k = A_t^y F\left(k_t, n_t^y\right),$$

and the construction sector satisfies

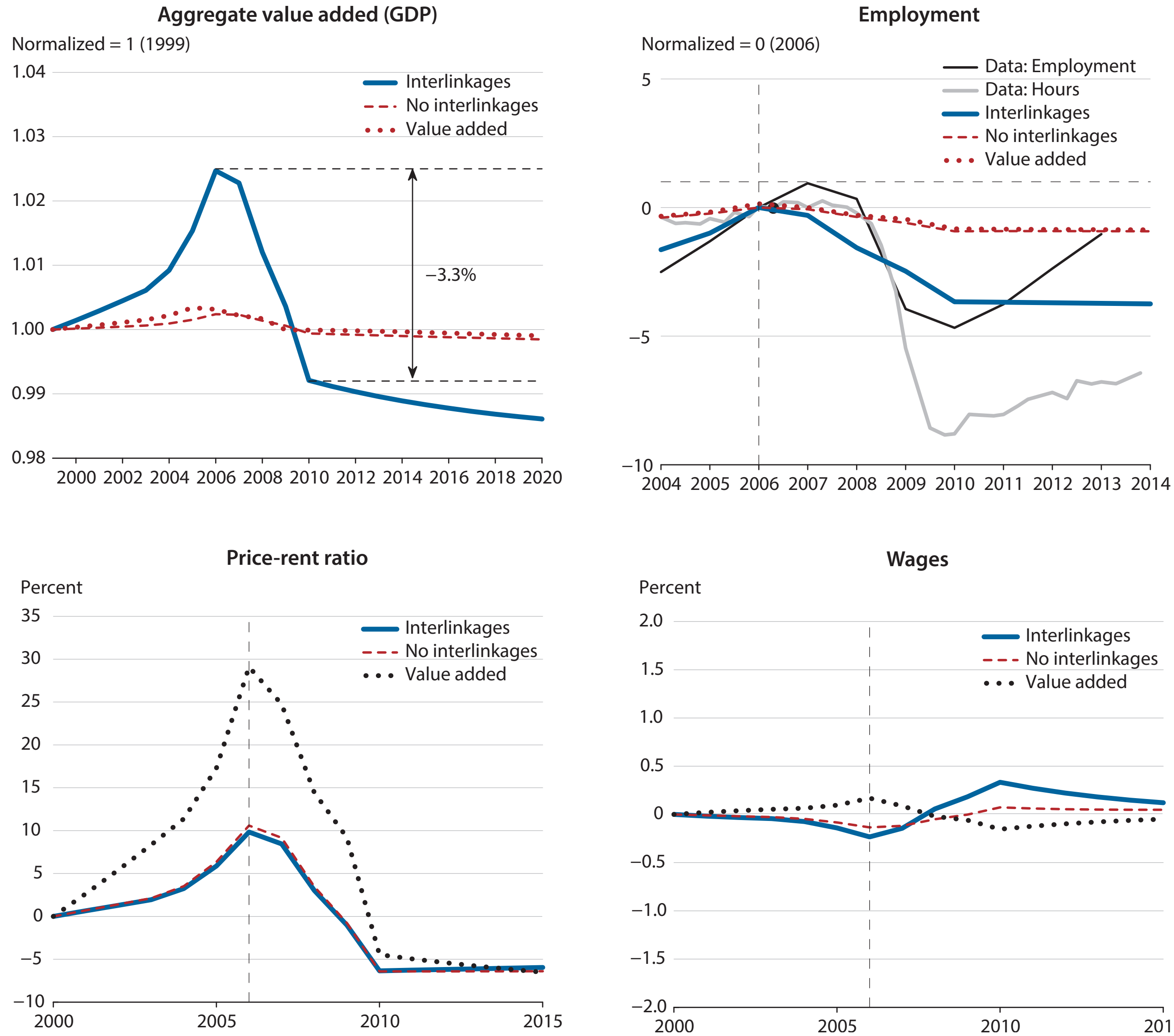
$$x_t^h = A_t^s G\left(n_t^s\right).$$

For both specifications, we carry out the same simulation experiment under the assumption that $\rho = 5$. The housing demand shifter is adjusted to generate movements in construction employment consistent with the data. Figure 12 compares the key macroeconomic aggregates in the three cases: the baseline model with interlinkages, a specification with no interlinkages, and the value-added specification.

Consider first the case labeled "no interlinkages," which is simpler because the intermediates are fixed to the initial steady-state levels. Both sectors are committed to producing the same amount of intermediates every period. During the housing boom, the only way to produce more structures is to use more capital and labor. Since the quantity of intermediates cannot adjust, prices adjust more relative to the baseline level. Qualitatively speaking, the equilibrium dynamics of this version of the model are similar to those of the baseline one. However, the quantitative implications are very different. Since intermediates are constant, the marginal product of labor in the construction sector does not increase as much as in the baseline experiment and employment also does not increase as much either. The construction sector expands during the boom, but because the links to the other sector have been severed,

Figure 12

The Impact of Construction: Role of Interlinkages (ρ = 5)

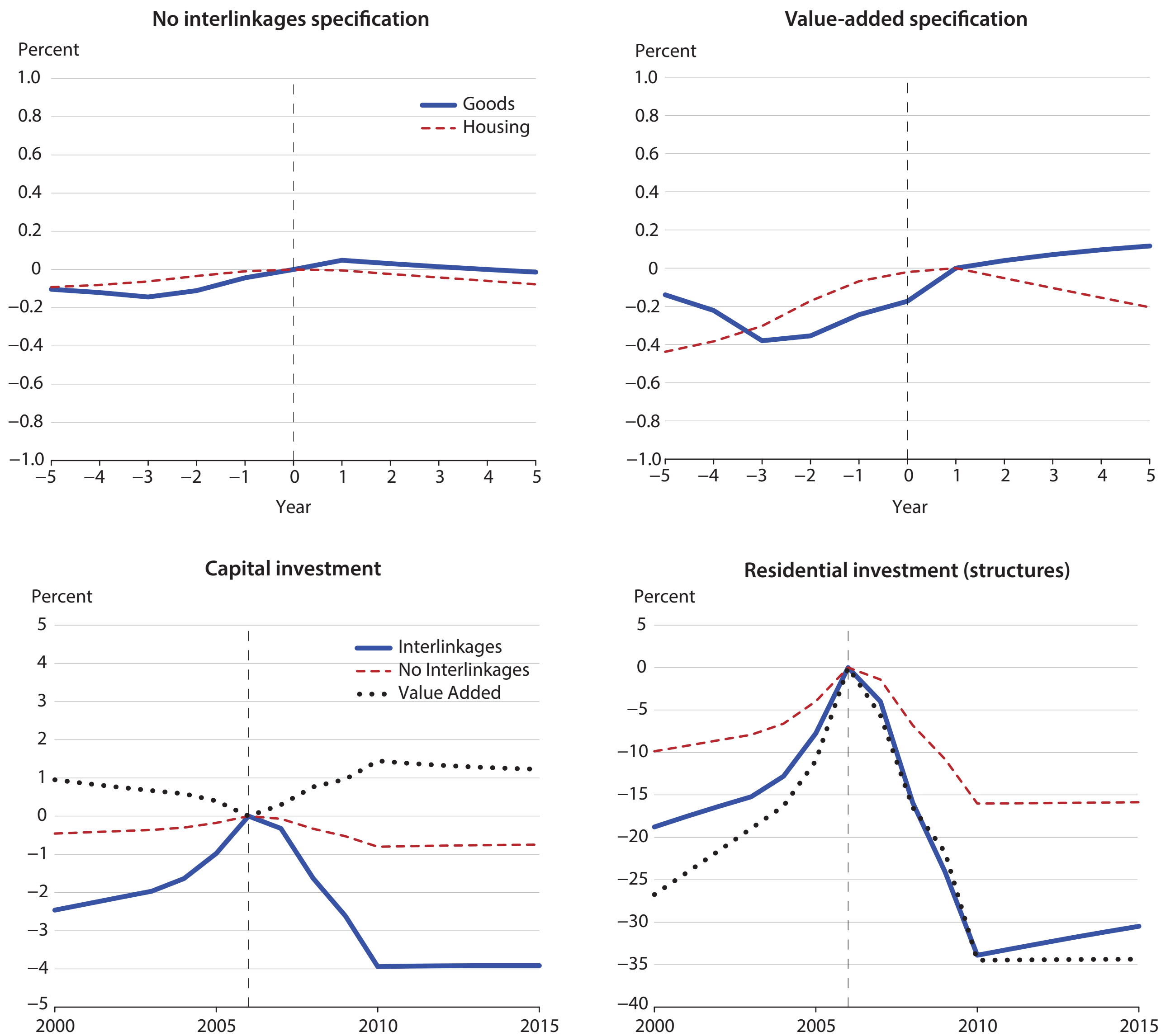


SOURCE: BEA and authors' calculations.

the latter barely moves in spite of the consumption complementarity: All movements are less than 0.5 percent. Consequently, the changes in GDP and employment are an order of magnitude smaller than in the economy with intersectoral links. The input-output links operate, de facto, as total factor productivity changes in the manufacturing sector, turning the variations in the demand for houses into a variation in the marginal value of output for the second sector.

Figure 13

Lead-Lag Responses of Consumption, Housing Spending, and Investment



SOURCE: BEA and authors' calculations.

In the value-added model, the change in the demand for housing also generates very small booms and busts in output and employment. Here the propagation from housing to the rest of the economy travels only on the demand side (via consumption complementarity), and the effect is consequently small. For the value-added case, the adjustment via prices is more severe because the only way to prevent individuals from purchasing more housing is to increase house prices. In this case, the effects of a positive demand shock are a sizeable appreciation of

Table 3

Quantitative Implications of Alternative Models

	Share of changes accounted for by the construction sector (percent)			
	Expansion 2000-07		Recession 2007-10	
Experiment	Employment	GDP	Employment	GDP
Baseline ($\rho = 5$)	60.2	25.3	43.9	56.2
Lower complementarity ($\rho = 3$)	28.7	14.1	28.7	40.6
Value-added specific ($\rho = 5$)	14.9	3.2	14.5	8.5
No interlinkages specific ($\rho = 5$)	14.5	2.5	10.6	4.8

SOURCE: Authors' calculations.

house values but small macroeconomic spillovers on the production side, as the construction sector is not directly interconnected with the rest of the economy. The adjustment in relative prices leads wages to remain relatively constant.

As shown in Figure 13, the responses of consumption, house spending, and investment are very different in each specification. In the model with fixed interlinkages, the dynamics are similar to those in the baseline model presented in Figure 9. However, the magnitudes are significantly smaller. The dynamics of the value-added specification are very different and resemble the case of high elasticity of substitution.

The study of these three alternative specifications illustrates an important point. The presence of interlinkages is *necessary* to generate large aggregate changes from fluctuations in construction. In fact, *both* alternative models generate very small changes in output and employment (for given shifts in the demand for housing) even though both maintain the complementarity between consumption and housing. Complementarity between housing and consumption, alone, delivers only very small aggregate fluctuations, which instead appear when the input-output structure of the economy is accounted for. Interlinkages are also crucial for the behavior of investment. In response to demand shocks, the model with interlinkages generates a simultaneous increase in consumption (housing and nonhousing) as well as investment (residential and nonresidential), whereas the value-added specification fails to account for such strong comovements. No asymmetric input-output structure, no business cycle action.

5.5 Quantitative Implications of Alternative Models

The different specifications studied lead to a general conclusion: The aggregate importance of the construction sector is significant despite its relatively small share in terms of employment and value added. Table 3 presents a summary of all the results discussed above. The table shows, for each of the model specifications considered, the fractions of the changes in employment and GDP accounted for by shocks to the construction sector during the expansion (2000-07) and the recession (2007-10). In light of the previous discussion, the numerical values should be easy to interpret at this point. The left side of Table 3 considers the role of the construction sector in the expansion. Regardless of the complementarity between housing and consumption goods, the model with interlinkages reveals that the construction sector accounts for a very significant share of the growth in total employment: between 29 and 60 percent. It also reveals that the contribution of construction to GDP is also much larger than its share: between 14 and 25 percent, which is somewhat smaller than its contribution to employment.

According to our model, the contribution of construction to employment and output was arguably even more important during the Great Recession. Depending on the specification, the decline in employment generated in the models with interlinkages is between 28 and 44 percent of the actual decline during the recession. In the case of GDP, the model generates between 41 and 56 percent of the observed changes. The model suggests that construction has been an important macroeconomic driver during the housing boom and bust and also highlights the asymmetry of its contribution between the expansion and the recession. During expansions the spillover on employment is larger than on output, but during recessions it is the opposite.

6 INTERLINKAGES AND BUSINESS CYCLE ACCOUNTING

An alternative methodology to identify the sources of economic fluctuations, within the context of a one-sector growth model, is "business cycle accounting," based on Chari, Kehoe, and McGrattan (2007). Recent works, including Arellano, Bai, and Kehoe (2019) and Ohanian and Raffo (2012), document that the Great Recession can be accounted for, mostly, by a worsening of labor market distortions. Both of these studies find that the labor wedge worsened by about 12 percent during the 2007-09 recession. Different explanations have been proposed to rationalize the measured increase in distortions in the labor market. For instance, Arellano, Bai, and Kehoe (2019) propose a model of imperfect financial markets and firm-level volatility. Such a model captures about half of the worsening in the labor wedge.

The wedge can be computed using data on employment, consumption, and wages generated by any model. It is defined as

$$X_t = -\frac{U_{N_t}}{U_{C_t}} / w_t,$$

where U_{N_t} is the marginal disutility measured at the aggregate level of employment, U_{C_t} is the marginal utility of consumption measured at the aggregate level of consumption, and w_t is the aggregate wage rate. Assuming wages are flexible and considering an aggregate Cobb-Douglas production function with capital share α, the wage can be replaced with

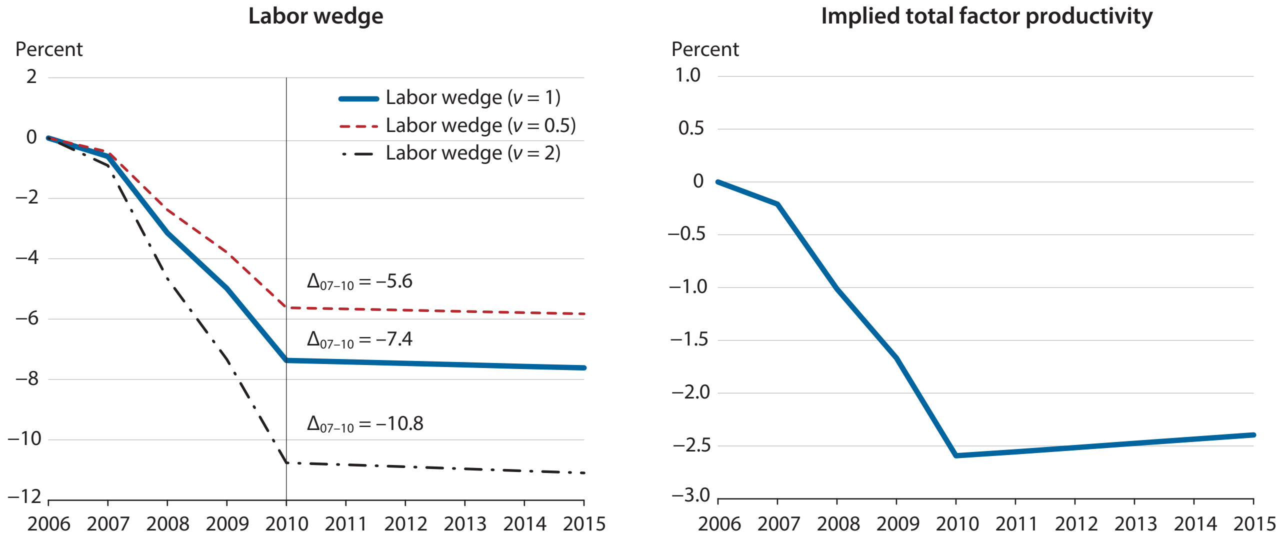
$$w_t = \frac{Y_t}{N}(1-\alpha).$$

Furthermore, using a log utility function for consumption and the following function for the disutility of employment,

$$U(N) = B\frac{N}{1+\upsilon}^{1+\upsilon},$$

Figure 14

Business Cycle Accounting of the Model with Interlinkages



SOURCE: BEA and authors' calculations.

the wedge can be written as

$$\Gamma_t = -\frac{B}{(1-\alpha)}\frac{C_t}{Y_t}N_t^{1+\upsilon}.$$

Notice that the parameters B and α are not important to understand fluctuations in the labor wedge; only the time series for aggregate consumption, output, and employment, and a value for υ, are required. We consider three values of $\upsilon = \{0.5, 1, 2\}$ and compute the labor wedge implied by our model using simulated data for consumption, output, and employment. Since our model has multiple sectors, several adjustments in the data are necessary. Consumption of goods and housing services are aggregated using relative prices $C_t = c_t + R_t h_t$. Aggregate output is $Y_t = C_t + X_t^k$, and total employment is $N_t = n_t^y + n_t^s$.

In the context of our model, any action in terms of implied distortions must be derived from the input-output structure and changes in relative prices. Figure 14 displays the changes in the labor wedge for our benchmark simulation. Its behavior is consistent with the data: The labor wedge worsens during the recession and does not recover quickly. For the case computed with $\upsilon = 1$, which is consistent with the value used by Ohanian and Raffo (2012), the labor wedge worsens by 7.4 percent; this is about 62 percent of the total change in the labor wedge during this period. Notice that our computation of the labor wedge assumes that wages are perfectly flexible. If this condition does not hold, the labor wedge has another component, referred to as the "firm-side" labor wedge in Arellano, Bai, and Kehoe (2019).[16] This wedge is basically the difference between the marginal product of labor and the wage. These authors refer to the other component of the labor wedge as the "consumer-side" labor wage, which is basically our Γ_t. Arellano, Bai, and Kehoe (2019) find that (i) the firm-side labor wedge has been fairly flat since 2006 and (ii) a worsening of the consumer-side labor wedge accounts for most of the Great Recession. Recall that there are no frictions in our model, so wages equal the marginal product of labor in every period. Thus, not only the behavior of the labor wedge during the Great Recession but also its decomposition in the data are consistent with what our model predicts. It can also be shown that our model would be consistent with a large and fairly persistent negative shock to total factor productivity. The combination of these two would rationalize the model-predicted Great Recession.

7 CONCLUSIONS

This article analyzes the contribution of the construction sector to U.S. economic growth, particularly during the Great Recession, using a two-sector input-output model. Historically, the construction sector has been relatively small in terms of its contributions to employment and GDP, but it is highly interconnected with other sectors in the economy and highly volatile. Our empirical analysis reveals how these sectoral interlinkages propagate changes in housing demand, greatly amplifying their effect on the overall economy. In our model, construction accounts for 52 percent of the decline in employment and 35 percent of the decline in output during the Great Recession and for similar, albeit slightly smaller, shares during the preceding boom.

The importance of the sectoral interlinkages is illustrated first using a simple static multi-sector model. We prove that, in our model, changes in housing demand have a much larger effect on aggregate activity when the sectors are asymmetrically interconnected. Also, the presence of irreversibility constraints on investment introduces an asymmetry between the expansion and the recession in the dynamic model. The simulation exercise is calibrated to reproduce the boom-bust dynamics of construction employment in the period 2002-10. In the model, during the housing boom all sectors expand and contribute 2 percent and 2.5 percent to the growth of output and employment, respectively. During the housing bust, the irreversibility constraint binds, amplifying the asymmetric response: The declines in output and employment are 3.3 percent and 3.8 percent, respectively. With a lower degree of complementarity, the asymmetric effect is not as large but still significant. These numbers can be used to calculate the contribution of construction in the data. The model suggests that during the expansion (2002-07), the construction sector accounted for a significant share of the growth in employment (between 29 and 60 percent) and GDP (between 14 and 25 percent). The construction sector's contribution was more important during the Great Recession (2007-10): Our calibrated model suggests that movements in housing demand—propagating through the economy—accounted for 29 to 44 percent of the variation in aggregate employment and 41 to 56 percent of the variation in GDP.

The presence of intersectoral linkages substantially amplifies the impact of changes in housing demand. In the model specifications without this mechanism, changes in housing demand consistent with the dynamics in construction employment have only a small effect on macroeconomic quantities. This is true even when the complementarity between consumption goods and housing services is high. A direct implication of this result is that the presence of interlinkages is *necessary* to generate large aggregate variations from changes in construction, and a high degree of complementarity is not *sufficient* to obtain the propagation of adjustments in housing demand to the rest of the economy we obtained in our model. To capture the intricacies of this mechanism it is necessary to formalize the aggregate economy with a multisector model with asymmetric interlinkages.

Since in our model the equilibrium is efficient, the behavior of output is also the behavior of potential output. Taking into account that both output and potential output were affected during the Great Recession, we perform a business cycle accounting exercise on simulated data from the model using the now common "wedge" approach. Despite the lack of any friction or distortion, the data our model generates attribute the recession to a worsening in the labor wedge. The magnitude generated by the model accounts for 62 percent of the total change observed in the data. Clearly, in our model, the metrics of the wedge measures are explained by the fact that we account for sectoral linkages, irreversibilities, and the movements of relative prices between sectors; they are not explained by frictions. This approach shows how multi-sector models of the business cycle can improve, or at least challenge, our understanding of the factors driving aggregate fluctuations.

A direct policy implication of our findings is that the output gap could be lower than historical estimates suggest. The historical anomalies in the events that took place between 2007 and 2013 can be accounted for by the equally anomalous evolution of housing demand in the six years before 2007 and in those following it. As far as policy is concerned, the basic implication of our research is simple: Estimations of output gaps using pre-2007 trends, and aggregate one-sector models, may lead to misleading policy prescriptions. ■

APPENDIX

A1 MICROFOUNDATIONS FOR THE HOUSING DEMAND SHIFTERS

The modeling strategy used in the article uses changes in "effective" housing demand as the driver of the amplification mechanism that affects sectoral interlinkages. There are many potential drivers of the changes in housing demand during the period we study. This particular episode witnessed sizeable changes in home ownership and significant innovations in housing finance at the household level (i.e., new mortgage products) and the industry level (i.e., the use of mortgage-backed securities as a liquid asset). This section provides a microfoundation of housing demand shocks using two different specifications. The first one uses credit constraints in housing finance, where changes in collateral requirements (i.e., loan-to-value ratios) are isomorphic to variations in the relative weight that housing has on preferences (intensive margin). The second one considers the case where a large number of new households enter into the owner-occupied housing market. At the aggregate level this is also captured as an increase in the aggregate demand for housing (extensive margin). Either specification would be consistent with the approach used by Chari, Kehoe, and McGrattan (2007) that reduces all the frictions in the model to distortions/wedges in the equilibrium conditions.

A1.1 "Effective" Housing Demand and Credit Constraints

The first specification relates housing demand to the presence of credit constraints. Consider a simple two-period version of the household optimization problem to allow for borrowing and collateral constraints of the form $b \leq \lambda ph$. This is the standard constraint that restricts the amount of housing finance to be proportional to the value of the house, $\lambda \in [0,1]$. The cost of borrowing, $R > 1$, is paid in the second period and can differ from the return of other assets,[17] r. For ease of exposition, consider the case where housing fully depreciates at the end of the second period and labor is inelastically supplied. The optimization problem of the representative consumer is

$$\max_{c_1, c_2, h, s, b} u(c_1) + \beta u(c_2) + \tilde{\theta} v(h),$$
$$s.t. \quad c_1 + s + ph = w_1 + b,$$
$$b \leq \lambda ph,$$
$$c_2 = w_2 + (1+r)s + bR.$$

The optimality condition for housing measured in terms of $t = 1$ consumption goods can be written as[18]

$$\theta \frac{v'(h)}{u'(c_1)} = p\left[1 - \lambda(1+\phi)\right].$$

When the solution is interior, increases in the value of λ reduce the cost of housing relative to the cost of consumption goods. This is observationally equivalent to an exogenous increase in θ. Similarly, a tightening of credit conditions reduces housing demand. From this perspective, the relevant value is $\tilde{\theta} = \theta / \left[1 - \lambda(1+\phi)\right]$. When housing finance is not present, $\lambda = 0$, the expression for housing demand is the same as in the previous model and $\tilde{\theta} = \theta$.

A1.2 "Effective" Housing Demand and Home Ownership

Part of the housing boom was fueled by an increase in the home ownership rate.[19] The second specification relates the aggregate change in housing demand to increasing participation in the housing market, using a model based on Garriga, Manuelli, and Peralta-Alva (2019). Consider an economy where households are ex-ante heterogeneous in their labor ability $\varepsilon \in \left[\underline{\varepsilon}, \bar{\varepsilon}\right]$ and the ability distribution is uniform $\varepsilon \sim U(\underline{\varepsilon}, \bar{\varepsilon}) \equiv f(\varepsilon)$. Preferences are represented by a utility function $u(c,h) = c(\theta + h)$, where consumption goods are perfectly divisible, $c \in R_+$, and housing is a discrete good with only one size of home available, $h \in \{0, \bar{h}\}$. The implicit assumption is that renters consume zero housing and homeowners consume a fixed positive amount; we could allow for the purchase of different size homes at the cost of introducing unnecessary notation. The parameter $\theta > 0$ can be interpreted as a reservation value for rental housing, and as $\theta \to 0$ owner-occupied housing becomes more desirable.

The optimization problem for the consumer is

$$
\begin{aligned}
v(\varepsilon) = \max_{h} & \left\{ u^r\left(c^r, 0\right), u^o\left(c^o, \bar{h}\right) \right\}, \\
s.t. \quad & c^o = w\varepsilon - \left(p\bar{h} + \phi\right), \\
& c^r = w\varepsilon,
\end{aligned}
$$

where w is the wage, p is the house price, and the price of consumption goods is 1. The term ϕ represents an exogenous transaction cost associated with buying a house, measured in terms of consumption goods. The optimal decision rule determines a cut-off level of ability necessary to purchase a house. For the specified preferences and under the necessary assumptions for an interior solution, the threshold for homeownership, ε^*, is characterized by

$$
\varepsilon^*\left(\theta, \bar{h}, \phi, p, w\right) \geq \frac{p}{w}\left(\theta + \bar{h}\right) + \frac{\phi}{w\bar{h}}.
$$

In the model, the determinants of ownership are the cost of housing relative to income, p/w; the house size, $\bar{h}$; transaction costs, ϕ; and the reservation value of rental housing,[20] θ. The comparative statics are straightforward. Increases in the house price, minimum house size, and transaction costs increase the income threshold required for home ownership, whereas an increase in wage income decreases it. Notice that the demand shifter changes the number of individuals buying houses, as the size of the latter is fixed.

Given this threshold, aggregate housing demand and the home ownership rate are in fact proportional:

$$
H\left(p\right) = \bar{h} \int_{\varepsilon^*}^{\bar{\varepsilon}} U\left(\underline{\varepsilon}, \bar{\varepsilon}\right) d\varepsilon = \frac{\bar{h}}{\left(\bar{\varepsilon} - \underline{\varepsilon}\right)} \left[\bar{\varepsilon} - \frac{p}{w}\left(\theta + \bar{h}\right) - \frac{\phi}{w\bar{h}} \right].
$$

Despite its simplicity, the expression shows the connection between housing demand and the key individual variables. A reduction in the rental threshold, θ, affects the total quantity demanded from the construction sector but also reduces the transaction costs, ϕ, affecting housing demand.

A2 ALTERNATIVE SPECIFICATIONS QUANTITATIVE MODEL: FIXED INTERLINKAGES AND VALUE-ADDED ECONOMIES

In our quantitative analysis, we made an effort to disentangle the role of interlinkages. It is always challenging to compare different models, but the quantitative analysis suggests similar results from the various alternatives. The first alternative considers an economy calibrated to the same initial steady state (parameters and targets) and compares the economy with interlinked production with an economy where the amount of intermediates is fixed at the initial steady-state level. The second alternative compares the value-added economy with the linkage economy. Both are calibrated to the same target values for the baseline year, but the underlying parameters are different.

For a given sequence of land $\left\{\bar{l}_t\right\}_{t=0}^{\infty}$, there is an optimization problem that solves for the equilibrium in each case. In the baseline case with interlinkages, the social planner chooses a sequence of quantities $\left\{c_t, x_t^k, x_t^s, n_t^y, n_t^s, m_t^{s,s}, m_t^{y,s}, m_t^{y,y}, m_t^{s,y}\right\}_{t=0}^{\infty}$ to maximize

$$\max \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t, \theta_t, h_t\right) + \gamma v\left(1 - n_t^y - n_t^s\right) \right],$$

$$\begin{aligned}
s.t. \quad & c_t + x_t^k + m_t^{y,y} + m_t^{y,s} = A_t^y F\left(k_t, n_t^y, m_t^{y,y}, m_t^{s,y}\right), \quad \forall t, \\
& x_t^s + m_t^{s,s} + m_t^{s,y} = A_t^s G\left(n_t^s, m_t^s\left(m_t^{s,s}, m_t^{y,s}\right)\right), \quad \forall t, \\
& x_t^k = k_{t+1} - \left(1 - \delta_k\right) k_t \geq 0, \quad \forall t, \\
& x_t^s = s_{t+1} - \left(1 - \delta_s\right) s_t \geq 0, \quad \forall t, \\
& h_t = H\left(s_t, \bar{l}_t\right), \quad \forall t, \\
& s_0, k_0 \geq 0.
\end{aligned}$$

In the model with no interlinkages, the production of intermediate goods is fixed at the steady-state level before the boom. In this case, the social planner is forced to produce the same quantity of intermediates each period $\left(m_t^{s,s} = m_0^{s,s}, m_t^{y,s} = m_0^{y,s}, m_t^{y,y} = m_0^{y,y}, m_t^{s,y} = m_0^{s,y}\right)$. To satisfy this constraint, the social planner picks a vector $\left\{c_t, x_t^k, x_t^s, n_t^y, n_t^s\right\}_{t=0}^{\infty}$ to maximize

$$\max \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t, \theta_t h_t\right) + \gamma v\left(1 - n_t^y - n_t^s\right) \right],$$

$$\begin{aligned}
s.t. \quad & c_t + x_t^k = A_t^y F\left(k_t, n_t^y, m_0^{y,y}, m_0^{s,y}\right) - \left(m_0^{y,y} + m_0^{y,s}\right), \quad \forall t, \\
& x_t^s = A_t^s G\left(n_t^s, m_0^s\left(m_0^{s,s}, m_0^{y,s}\right)\right) - \left(m_0^{s,s} + m_0^{s,y}\right), \quad \forall t, \\
& x_t^k = k_{t+1} - \left(1 - \delta_k\right) k_t \geq 0, \quad \forall t, \\
& x_t^s = s_{t+1} - \left(1 - \delta_s\right) s_t \geq 0, \quad \forall t, \\
& h_t = H\left(s_t, \bar{l}_t\right), \quad \forall t, \\
& s_0, k_0, m_0^{s,s}, m_0^{y,s}, m_0^{y,y}, m_0^{s,y} \geq 0.
\end{aligned}$$

The last case studied is that of a value-added economy, where intermediate goods are completely eliminated. The social planner chooses a vector of quantities $\left\{c_t, x_t^k, x_t^s, n_t^y, n_t^s\right\}_{t=0}^{\infty}$ to maximize

$$\max \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t, \theta_t h_t\right) + \gamma v\left(1 - n_t^y - n_t^s\right) \right],$$

$$\begin{aligned}
s.t. \quad & c_t + x_t^k = A_t^y F\left(k_t, n_t^y\right), \quad \forall t, \\
& x_t^s = A_t^s G\left(n_t^s\right), \quad \forall t, \\
& x_t^k = k_{t+1} - \left(1 - \delta_k\right) k_t \geq 0, \quad \forall t, \\
& x_t^s = s_{t+1} - \left(1 - \delta_s\right) s_t \geq 0, \quad \forall t, \\
& h_t = H\left(s_t, \bar{l}_t\right), \quad \forall t, \\
& s_0, k_0 \geq 0.
\end{aligned}$$

A3 CALIBRATION OF INTERLINKAGES

Interlinkages are calibrated using input-output data. In particular, the information shown in Table A1 is used to calibrate the parameters of the two production functions and is calculated from the BEA's 2010 use input-output table. The use table shows the uses of commodities by intermediate and final users; rows present the commodities or products, and columns display the industries and the final users. The sum of the entries in a row is the output of that commodity. The columns show the products consumed by each industry and the three components of value added—compensation of employees, taxes on production and imports less subsidies, and the gross operating surplus. Value added is the difference between an industry's output and the cost of its intermediate inputs; total value added is equal to GDP.

Table A1 displays input-output values (which are originally in millions of dollars) as a fraction of each industry's output. Construction receives most of its inputs from other industries (48.3 percent of its gross output) and less than 1 percent from itself. The reverse is true for the other industries, as they receive most of their inputs from themselves (43.0 percent of their total gross output).

Table A1

Coefficients with Respect to Column Industries

Commodities/industries	Construction	Other industries
Construction	0.0009	0.0058
Other industries	0.4828	0.4301
Compensation of employees	0.3625	0.2802
Taxes on production and imports, less subsidies	0.0072	0.0471
Gross operating surplus	0.1466	0.2368
Total	1.0000	1.0000

SOURCE: BEA 2010 use input-output table.

NOTES

1 Davis and Heathcote (2005) construct a real business cycle with housing and interlinkages. In the baseline economy with Cobb-Douglas preferences, the presence of interlinkages generates a relatively small contribution to aggregate fluctuations in response to productivity shocks. Our theoretical model shows that Cobb-Douglas preferences completely eliminate the role of interlinkages because of insufficient complementarity. This is true even when the sectoral linkages are asymmetric. Iacoviello (2005) generates house price fluctuations using shocks to Cobb-Douglas preferences, but the productive structure of the economy does not have interlinkages, so most of the action is driven by the presence of binding collateral constraints and price rigidities.

2 The appendix provides two microfoundations for the drivers of housing demand. We show that this shock is isomorphic to a model that relaxes collateral constraints or a model with housing demand at the extensive margin.

3 The analysis abstracts from both the increase in the burden of debt brought about by the decline in home prices (which is the focus of Garriga, Manuelli, and Peralta-Alva, 2019) and the reduction in credit activity it implied, two factors that are likely to have played a major role in the overall process. Although these factors could interact with

the sectoral interlinkages, abstracting from them captures the contribution of the real side of the economy in the recession.

4 It is often argued that the housing sector is of great relevance to the aggregate economy because housing wealth is a major determinant of consumption demand (see Carroll, Otsuka, and Slacalek, 2011; Case, Quigley, and Shiller, 2005; and Mishkin, 2007, which are among the most cited articles). An ample and somewhat more recent literature (e.g., Calomiris, Longhofer, and Miles, 2009, and Iacoviello, 2011, and the references therein) has cast serious doubts on the quantitative relevance of this channel for business cycles analysis. While the housing sector is certainly very cyclical, this is most likely not due to a causal chain going from housing wealth to consumption and aggregate demand to output, but to a host of other common factors driving such comovements. Further, the same literature also reveals that, when empirical evidence of a causal link is found, the latter is not only quantitatively weak but its magnitude is also dependent on demographic and financial variables.

5 The traditional view of the business cycle literature is that idiosyncratic sectoral shocks are likely to average out and have no aggregate effects as the number of sectors in the economy gets larger (i.e., Lucas, 1981; Kydland and Prescott, 1982; Long and Plosser, 1983; and Dupor, 1999).

6 In the analysis hereafter, the definition of the construction sector does not include the real estate and leasing sector, because the two sectors are quite different. Including it in the definition of the construction sector would substantially increase the construction sector's significance in accounting for the Great Recession.

7 For employment, we use the employment requirements matrix from the Bureau of Labor Statistics (BLS).

8 For expositional purposes, in this section we assume that the diagonal coefficients of the requirement matrix A in the Leontief model are zero. The general formulation would be $c+m_{yy}+m_{yh}=Y=A_y f\left(n_y,\varepsilon_{yy}m_{yy},\varepsilon_{yh}m_{hy}\right)$ and $h+m_{hy}+m_{hh}=H=A_h g\left(n_h,\varepsilon_{hy}m_{yh},\varepsilon_{hh}m_{hh}\right)$, where in each m_{ij}, the first subscript denotes the origin (i) and the second denotes the destination (j).

9 As $\varepsilon_j \to 0$, the required quantity of the intermediate good converges to zero, $m_j \to 0$. When both coefficients converge to zero, the technologies become $c = A_y n_y$ and $h = A_h n_h$, respectively. In this case, the interlinkages disappear.

10 This condition is the well-known "all or nothing" property of Leontief input-output models. When it is met, the economy is productive and any nonnegative value added is reachable if enough labor input is available.

11 This condition is related to the irrelevance result in Dupor (1999).

12 This specification implies a Frisch elasticity of labor equal to 2. Keane and Rogerson (2012) argue that this elasticity can be reconciled with lower elasticity estimates at the micro level.

13 The long-run value of θ has small quantitative implications for the short-run dynamics discussed in this article.

14 The magnitudes of these numbers vary with the time interval considered, but the overall magnitudes are within reasonable bounds.

15 See the appendix for model details.

16 They follow Galí, Gertler, and López-Salido (2007) in this decomposition.

17 In this class of model, the consumer usually has an incentive to borrow to purchase the house when the cost of borrowing is lower than the return of other assets, $R/1 + r = \phi < 1$ (additional conditions are discussed later in the article). In many countries, interest payments are tax deductible, reducing the effective cost of borrowing relative to other assets. Under this assumption, the Lagrange multiplier of the collateral constraint binds, and housing demand directly determines the amount of borrowing.

18 For an interior solution with borrowing it suffices that $\lambda < 1/(1 + \phi)$.

19 See Chambers, Garriga, and Schlagenhauf (2009a,b) for a detailed discussion on the home ownership rate boom between 1994 and 2007.

20 When the transaction cost is proportional to the value of the house, the budget constraint of the buyer is slightly different, $c^o = w\varepsilon-\left(p+\phi\right)\bar{h}$, and the homeownership threshold is $\varepsilon^* \geq \frac{\left(p+\phi\right)}{w}\left(\theta+\bar{h}\right)$.

REFERENCES

Acemoglu, Daron; Carvalho, Vasco M.; Ozdaglar, Asuman and Tahbaz-Salehi, Alireza. "The Network Origins of Aggregate Fluctuations." *Econometrica*, 2012, *80*(5), pp. 1977-2016; https://doi.org/10.3982/ECTA9623.

Acemoglu, Daron; Ozdaglar, Asuman and Tahbaz-Salehi, Alireza. "Systemic Risk and Stability in Financial Networks." *American Economic Review*, 2015, *105*(2), pp. 564-608; https://doi.org/10.1257/aer.20130456.

Arellano, Cristina; Bai, Yan and Kehoe, Patrick J. "Financial Frictions and Fluctuations in Volatility." *Journal of Political Economy*, October 2019, *127*(5), pp. 2049-103; https://doi.org/10.1086/701792.

Atalay, Enghin; Dratzburg, Thorsten and Wan, Zhenting. "Accounting for the Sources of Macroeconomic Tail Risks." *Economics Letters*, April 2018, *165*, pp. 65-69; https://doi.org/10.1016/j.econlet.2018.01.032.

Benhabib, J. and Nishimura, K. "On the Uniqueness of Steady States in an Economy with Heterogeneous Capital Goods." *International Economic Review*, 1979, *20*(1), pp. 59-82; https://doi.org/10.2307/2526416.

Black, Fischer. "Interest Rates as Options." *Journal of Finance*, 1995, *50*(5), pp. 1371-76; https://doi.org/10.1111/j.1540-6261.1995.tb05182.x.

Bloom, N. "The Impact of Uncertainty Shocks." *Econometrica*, 2009, *77*(3), pp. 623-85; https://doi.org/10.3982/ECTA6248.

Boldrin, M. and Deneckere, R. "Sources of Complex Dynamics in Two-Sector Growth Models." *Journal of Economic Dynamics and Control*, 1990, *14*, pp. 627-53; https://doi.org/10.1016/0165-1889(90)90036-G.

Boldrin, M. and Horvath, Michael. "Labor Contracts and Business Cycles." *Journal of Political Economy*, October 1995, *103*(5), pp. 972-1004; https://doi.org/10.1086/262010.

Brunnermeier, Markus and Sannikov, Yuliy. "Reviving 'Money and Banking,'" in Lucrezia Reichlin and Richard Baldwin, eds., *Is Inflation Targeting Dead?* Centre for Economic Policy Research, 2013, pp. 95-101.

Caliendo, Lorenzo; Parro, Fernando; Rossi-Hansberg, Esteban and Sarte, Pierre-Daniel. "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy." NBER Working Paper 20168, National Bureau of Economic Research, 2014; https://doi.org/10.3386/w20168.

Calomiris, Charles; Longhofer, Stanley D. and Miles, William. "The (Mythical?) Housing Wealth Effect." NBER Working Paper 15075, National Bureau of Economic Research, 2009; https://doi.org/10.3386/w15075.

Campbell, Jeffrey R. and Hercowitz, Zvi. "The Role of Collateralized Household Debt in Macroeconomic Stabilization." NBER Working Papers 11330, National Bureau of Economic Research, 2005; https://doi.org/10.3386/w11330.

Carroll, Christopher D.; Otsuka, Misuzu and Slacalek, Jiri. "How Large Are Housing and Financial Wealth Effects? A New Approach." *Journal of Money, Credit, and Banking*, February 2011, *43*, pp. 55-79; https://doi.org/10.1111/j.1538-4616.2010.00365.x.

Carvalho, Vasco M. "Aggregate Fluctuations and the Network Structure of Intersectoral Trade." Economics Working Papers, Department of Economics and Business, Universitat Pompeu Fabra, 2010.

Carvalho, Vasco M. and Gabaix, Xavier. "The Great Diversification and Its Undoing." *American Economic Review*, 2013, *103*(5), pp. 1697-1727; https://doi.org/10.1257/aer.103.5.1697.

Case, Karl E.; Quigley, John M. and Shiller, Robert J. "Comparing Wealth Effects: The Stock Market Versus the Housing Market." *Advances in Macroeconomics*, *5*(1), 2005; https://doi.org/10.2202/1534-6013.1235.

Chambers, Matthew; Garriga, Carlos and Schlagenhauf, Don. "Accounting for Changes in the Homeownership Rate." *International Economic Review*, 2009a, *50*(3), pp. 677-726; https://doi.org/10.1111/j.1468-2354.2009.00544.x.

Chambers, Matthew S.; Garriga, Carlos and Schlagenhauf, Don. "The Loan Structure and Housing Tenure Decisions in an Equilibrium Model of Mortgage Choice." *Review of Economic Dynamics*, 2009b, *12*(3), pp. 444-68; https://doi.org/10.1016/j.red.2009.01.003.

Chari, V.V.; Kehoe, P.J. and McGrattan, E.R. "Business Cycle Accounting." *Econometrica*, 2007, *75*(3), pp. 781-836; https://doi.org/10.1111/j.1468-0262.2007.00768.x.

Christiano, Lawrence J.; Motto, Roberto and Rostagno, Massimo. "Financial Factors in Economic Fluctuations." European Central Bank working paper, May 2010.

Davis, Morris A. and Heathcote, Jonathan. "Housing and the Business Cycle." *International Economic Review*, 2005, *46*(3), pp. 751-84; https://doi.org/10.1111/j.1468-2354.2005.00345.x.

Davis, Morris A and Van Nieuwerburgh, Stijn. "Housing, Finance and the Macroeconomy." *Handbook of Urban and Regional Economics*, 2015, *5*, pp. 753-811; https://doi.org/10.1016/B978-0-444-59531-7.00012-0.

Dupor, Bill. "Aggregation and Irrelevance in Multi-Sector Models." *Journal of Monetary Economics*, April 1999, *43*(2), pp. 391-409; https://doi.org/10.1016/S0304-3932(98)00057-9.

Fisher, Jonas D.M. "Why Does Household Investment Lead Business Investment Over The Business Cycle?" *Journal of Political Economy*, 2007, *115*(1), pp. 141-68; https://doi.org/10.1086/511994.

Flavin, Marjorie and Nakagawa, Shinobu. "A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence." *American Economic Review*, March 2008, *98*(1), pp. 474-95; https://doi.org/10.1257/aer.98.1.474.

Foerster, Andrew T.; Sarte, Pierre-Daniel G. and Watson, Mark W. "Sectoral Versus Aggregate Ahocks: A Structural Factor Analysis of Industrial Production." *Journal of Political Economy*, 2011, *119*(1), pp. 1-38; https://doi.org/10.1086/659311.

Gabaix, Xavier. "The Granular Origins of Aggregate Fluctuations." *Econometrica*, 2011, *79*(3), pp. 733-72; https://doi.org/10.3982/ECTA8769.

Galí, J.; Gertler, M. and López-Salido, J.D. "Markups, Gaps, and the Welfare Costs of Business Fluctuations." *Review of Economics and Statistics*, 2007, *89*(1), pp. 44-59; https://doi.org/10.1162/rest.89.1.44.

Garriga, Carlos; Chambers, Matthew and Schlagenhauf, Don. "Did Housing Policies Cause the Postwar Boom in Homeownership?" NBER Working Paper 18821, National Bureau of Economic Research, 2012; https://doi.org/10.20955/wp.2012.021.

Garriga, Carlos; Manuelli, Rodolfo and Peralta-Alva, Adrian. "A Macroeconomic Model of Price Swings in the Housing Market." *American Economic Review*, 2019, *109*(6), pp. 2036-72; https://doi.org/10.1257/aer.20140193.

Gertler, Mark and Karadi, Peter. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics*, 2011, *58*(1), pp. 17-34; https://doi.org/10.1016/j.jmoneco.2010.10.004.

Gervais, Martin. "Housing Taxation and Capital Accumulation." *Journal of Monetary Economics*, 2002, *49*(7), pp. 1461-89; https://doi.org/10.1016/S0304-3932(02)00172-1.

Hall, R.E. "The Long Slump." *American Economic Review*, 2011, *101*, pp. 431-69; https://doi.org/10.1257/aer.101.2.431.

He, Zhiguo and Krishnamurthy, Arvind. "A Macroeconomic Framework for Quantifying Systemic Risk." *American Economic Journal: Macroeconomics*, October 2019, *11*(4), pp. 1-37; https://doi.org/10.1257/mac.20180011.

Herkenhoff, Kyle F. and Ohanian, Lee E. "Foreclosure Delay and U.S. Unemployment." Working Paper 2012-017, Federal Reserve Bank of St. Louis, 2012; https://doi.org/10.20955/wp.2012.017.

Herkenhoff, Kyle F. and Ohanian, Lee E. "Labor Market Dysfunction During the Great Recession." *Cato Papers on Public Policy*, July 2011, *1*; https://doi.org/10.3386/w17313.

Horvath, Michael. "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks." *Review of Economic Dynamics*, 1998, *1*(4), pp. 781-808; https://doi.org/10.1006/redy.1998.0028.

Horvath, Michael. "Sectoral Shocks and Aggregate Fluctuations." *Journal of Monetary Economics*, 2000, *45*(1), pp. 69-106; https://doi.org/10.1016/S0304-3932(99)00044-6.

Iacoviello, Matteo. "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle." *American Economic Review*, June 2005, *95*(3), pp. 739-64; https://doi.org/10.1257/0002828054201477.

Iacoviello, Matteo. "Housing Wealth and Consumption." International Finance Discussion Paper No. 1027, Board of Governors of the Federal Reserve System, 2011; https://doi.org/10.17016/IFDP.2011.1027.

Iacoviello, Matteo and Pavan, Marina. "Housing and Debt Over the Life Cycle and Over the Business Cycle." *Journal of Monetary Economics*, 2013, *60*(2), pp. 221-38; https://doi.org/10.1016/j.jmoneco.2012.10.020.

Jermann, Urban and Quadrini, Vincenzo. "Erratum: Macroeconomic Effects of Financial Shocks." *American Economic Review*, 2012, *102*(2), p. 1186; https://doi.org/10.1257/aer.102.2.1186.

Jones, Callum; Midrigan, Virgiliu and Philippon, Thomas. "Household Leverage and the Recession." NBER Working Paper 16965, National Bureau of Economic Research, June 2018.; https://doi.org/10.5089/9781484373866.001.

Keane, Michael and Rogerson, Richard. "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom." *Journal of Economic Literature*, June 2012, *50*(2), pp. 464-76; https://doi.org/10.1257/jel.50.2.464.

Kehoe, Timothy J.; Ruhl, Kim J. and Steinberg, Joseph B. "Global Imbalances and Structural Change in the United States." NBER Working Paper 9339, National Bureau of Economic Research, August 2013; https://doi.org/10.3386/w19339.

Kocherlakota, N. "Incomplete Labor Markets." Working paper, Federal Reserve Bank of Minneapolis, 2012; https://www.minneapolisfed.org/article/2012/incomplete-labor-markets.

Kydland, Finn E. and Prescott, Edward C. "Time to Build and Aggregate Fluctuations." *Econometrica*, 1982, pp. 1345-70; https://doi.org/10.2307/1913386.

Leamer, Edward E. "Housing Is the Business Cycle." NBER Working Papers 13428, National Bureau of Economic Research, 2007; https://doi.org/10.3386/w13428.

Li, Nan and Martin, Vance L. "Real Sectoral Spillovers: A Dynamic Factor Analysis of the Great Recession." *Journal of Monetary Economics*, November 2019, *107*, pp. 77-95; https://doi.org/10.1016/j.jmoneco.2018.10.002.

Long, John B. Jr. and Plosser, Charles I. "Real Business Cycles." *Journal of Political Economy*, 1983, pp. 39-69; https://doi.org/10.1086/261128.

Lucas, Robert E. *Studies in Business-Cycle Theory*. MIT Press, 1981.

Martinez, Leonardo; Hatchondo, Juan Carlos and Sánchez, Juan M. "Mortgage Defaults." Journal of Monetary Economics, 2015, *76*, pp. 173-90; https://doi.org/10.1016/j.jmoneco.2015.09.011.

Mishkin, Frederic S. "Housing and the Monetary Transmission Mechanism." NBER Working Paper 13518, National Bureau of Economic Research, 2007; https://doi.org/10.3386/w13518.

Mitman, Kur; Kaplan, Greg and Violante, Gianluca. "The Housing Boom and Bust: Model Meets Evidence." *Journal of Political Economy* (forthcoming).

Ohanian, L.E. and Raffo, A. "Aggregate Hours Worked in OECD Countries: New Measurement and Implications for Business Cycles." *Journal of Monetary Economics,* 2012, *59*(1), pp. 40-56; https://doi.org/10.1016/j.jmoneco.2011.11.005.

Rognlie, Matthew; Shleifer, Andrei and Simsek, Alp. "Investment Hangover and the Great Recession." *American Economic Journal: Macroeconomics*, 2018, *10*(2), pp. 113-53; https://doi.org/10.1257/mac.20160211.

The Case of the Reappearing Phillips Curve: A Discussion of Recent Findings

Asha Bharadwaj and Maximiliano Dvorkin

The Phillips curve seems to have flattened over time. In this article, we use a simple New Keynesian model to analyze potential pitfalls in the estimation of the slope of the structural Phillips curve. Changes in the conduct of monetary policy or in the relative importance of supply and demand shocks may bias simple estimations of the slope of the Phillips curve. Recent proposals have favored estimations using regional or city data in an effort to overcome these issues. We use a simple model of a monetary union with a continuum of economies and find that some of the drawbacks of the aggregate model are still present in a cross-section of many regions in a monetary union. The relative importance of the demand and supply shocks largely determines the empirical relation between unemployment and inflation in both the aggregate and the cross-section of regions. Our analysis shows potential pitfalls in estimating the slope of the Phillips curve, even if using regional data. (JEL E12, E31, E58, R13)

Federal Reserve Bank of St. Louis *Review*, Third Quarter 2020, *102*(3), pp. 313-37.
https://doi.org/10.20955/r.102.313-37

1 INTRODUCTION

Central banks around the world intervene in financial markets by setting the short-term nominal interest rates to stimulate economic activity and control inflation. There are many theories that suggest that changes in interest rates due to monetary policy have effects on real activity: An increase in interest rates is associated with a tightening of the economy and a decrease in the real output, while a decrease in interest rates is associated with an increase in real output. Interest rates can affect output through several channels. For instance, an increase in interest rates implies an increase in the return on savings. In this case, individuals have an incentive to save more and forego current consumption, which translates to a decrease in current real output, all else equal. Clearly, the opposite is true in the case of a decrease in interest rates. Other possible mechanisms through which movements in interest rates affect real activity include borrowing costs and investment. As interest rates decline it becomes cheaper to borrow and invest, so business investment goes up, increasing total output.

Asha Bharadwaj is a research associate and Maximiliano Dvorkin is a senior economist at the Federal Reserve Bank of St. Louis.

While these theories shed light on the effects of monetary policy on the real economy, there is no consensus on the link between interest rates and inflation. A widespread view among policymakers is that there is a trade-off between real activity and inflation and that monetary policy decisions on interest rates, by affecting the real economy, ultimately affect inflation. This trade-off is known as the Phillips curve.

The Phillips curve was popularized by A.W. Phillips in 1958, when he showed a statistically significant negative relation between the unemployment rate and the growth rate of nominal wages—that is, wage inflation. Based on this empirical relationship, Samuelson and Solow (1960) argued that a looser monetary policy could reduce the unemployment rate by allowing inflation to rise. This then implied that monetary authorities could exploit this trade-off. Since it was first discovered empirically, the Phillips curve has guided discussions of monetary policy and has shaped our understanding of the transmission of monetary policy to prices. More recently, several theories on price setting by firms can rationalize the existence of a Phillips curve in an economic model.

Over the years, the Phillips curve has received several criticisms. Recent articles have argued that inflation can be approximated by statistical processes unrelated to the amount of slack in the economy (Atkeson and Ohanian, 2001; Cecchetti et al., 2017; and Stock and Watson, 2007). Moreover, the lack of a stable relationship between inflation and various measures of slack has led several articles to conclude that the Phillips curve has weakened over the years (Blanchard, Cerutti, and Summers, 2015; and Coibion and Gorodnichenko, 2015).

Several articles have pushed back on this criticism and have attempted to "recover" the Phillips curve. Fitzgerald and Nicolini (2014) argue that aggregate data are uninformative about the true structural relationship between unemployment and inflation, and that in fact, under a specific definition of inflation targeting, the evolution of equilibrium inflation is a random walk. They then show that regional data can be used to identify the structural relationship between unemployment and inflation. The main intuition is that monetary policy typically reacts to the aggregate state of the economy, but not to regional conditions. Thus, it is possible to use the deviations of regional economic activity relative to the aggregate and the deviation of inflation relative to the aggregate to recover the relationship between unemployment and prices.

A recent article by McLeay and Tenreyro (2019) also supports this view. The authors argue that it is difficult to identify the slope of the Phillips curve empirically, even if a negative relationship does hold true in the underlying model. This is because monetary policy will react to economic shocks in order to stimulate output when it is below potential and reduce inflation when it is above target. The actions of the monetary authority will typically affect the empirical slope of the Phillips curve. McLeay and Tenreyro (2019) then use a simple New Keynesian model to highlight this estimation bias due to the endogeneity of monetary policy. They propose several solutions including using regional Phillips curves to circumvent this identification problem.

In our article, we follow the approach suggested by McLeay and Tenreyro (2019) closely and use a simple New Keynesian model to highlight the issues that arise with the identification of the empirical slope of the aggregate Phillips curve. We then use a simple New Keynesian model of a monetary union, as given by Gali and Monacelli (2008), to attempt to recover the Phillips curve at the regional level. However, we find that this approach is not sufficient to overcome the identification issues highlighted by McLeay and Tenreryo (2019), since even at the regional level our model fails to recover the slope of the Phillips curve. We argue that several factors, including the relative importance of the demand and cost-push shocks, affect the estimation of the slope of the Phillips curve, at both the aggregate and regional levels.

The structure of this article is as follows. Section 2 briefly reviews existing literature on the Phillips curve. Section 3 presents empirical evidence on the aggregate relationship between unemployment and inflation. Section 4 introduces a simple New Keynesian model of optimal policy with the Phillips curve and describes the empirical relationships we obtain when the model is used as a data-generating process. Section 5 discusses empirical challenges with regional Phillips curves. Section 6 presents a simple model of a monetary union with a continuum of regions and discusses the results we obtain when we use the model to simulate data. Section 7 concludes.

2 LITERATURE REVIEW AND BACKGROUND ON THE PHILLIPS CURVE

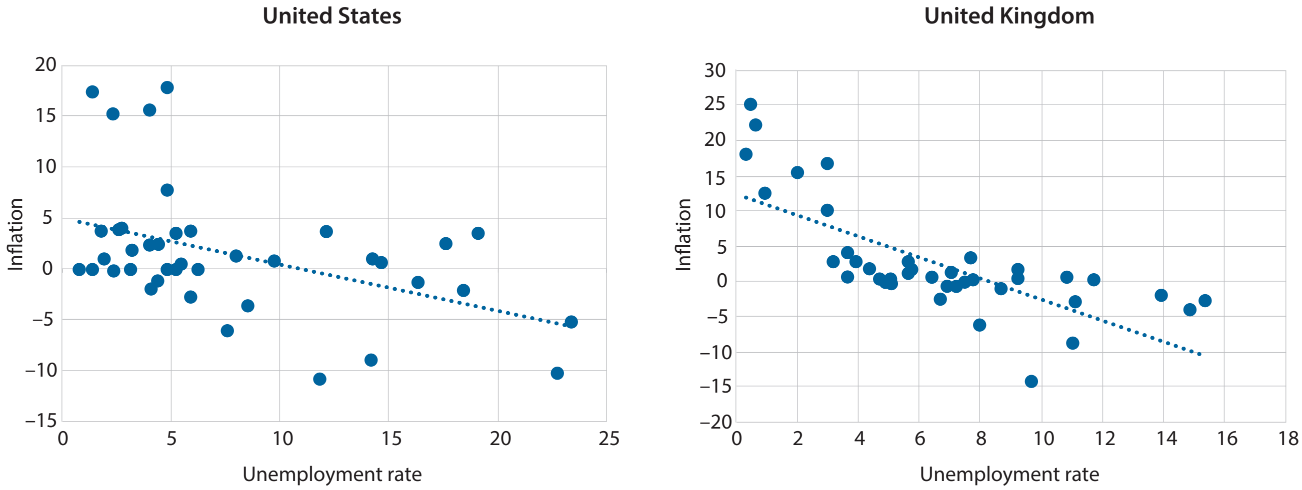
Phillips (1958) showed, for the United Kingdom and for the years 1861-1913, a statistically significant negative relationship between the unemployment rate and the growth rate of nominal wages held in the data. This result led to an outpouring of work on this topic, including the work by Samuelson and Solow (1960), who argued that policy could exploit the trade-off between inflation and the unemployment rate: Looser monetary policy would lead to an increase in inflation and a decrease in the unemployment rate. Figure 1 portrays the relationship between inflation and unemployment for the United States and United Kingdom, between 1900 and 1940, and we observe a clear negative trend between the level of unemployment and the growth rate of prices.

However, in the 1960s and 1970s, the relationship between unemployment and inflation started to change and was no longer a robust negative relationship. Figure 2 represents the relationship between unemployment and inflation in the 1960s for three countries—the United Kingdom, the United States, and France—and we observe that the negative relationship we would expect is no longer consistent across countries. It is negative for the United States, while it is positive for France and the United Kingdom. Lucas (1972) argued that the statistical relationship between unemployment and inflation depends on the parameters governing monetary policy and, therefore, is not "structural." Lucas provided a theoretical model consistent with the existence of the observed statistical relationship between inflation and unemployment, but in which systematic attempts by policy to exploit the trade-off are not successful and change the statistical relation of these two variables.

Gali and Gertler (1999) developed a structural model of the Phillips curve with microfoundations, based on initial work by Calvo (1983). They argued that real marginal costs and inflation expectations are significant determinants of inflation, as opposed to the backward-looking behavior that had been considered quantitatively important in the existing literature. Several articles then used this microfounded Phillips curve, also called the New Keynesian

Figure 1

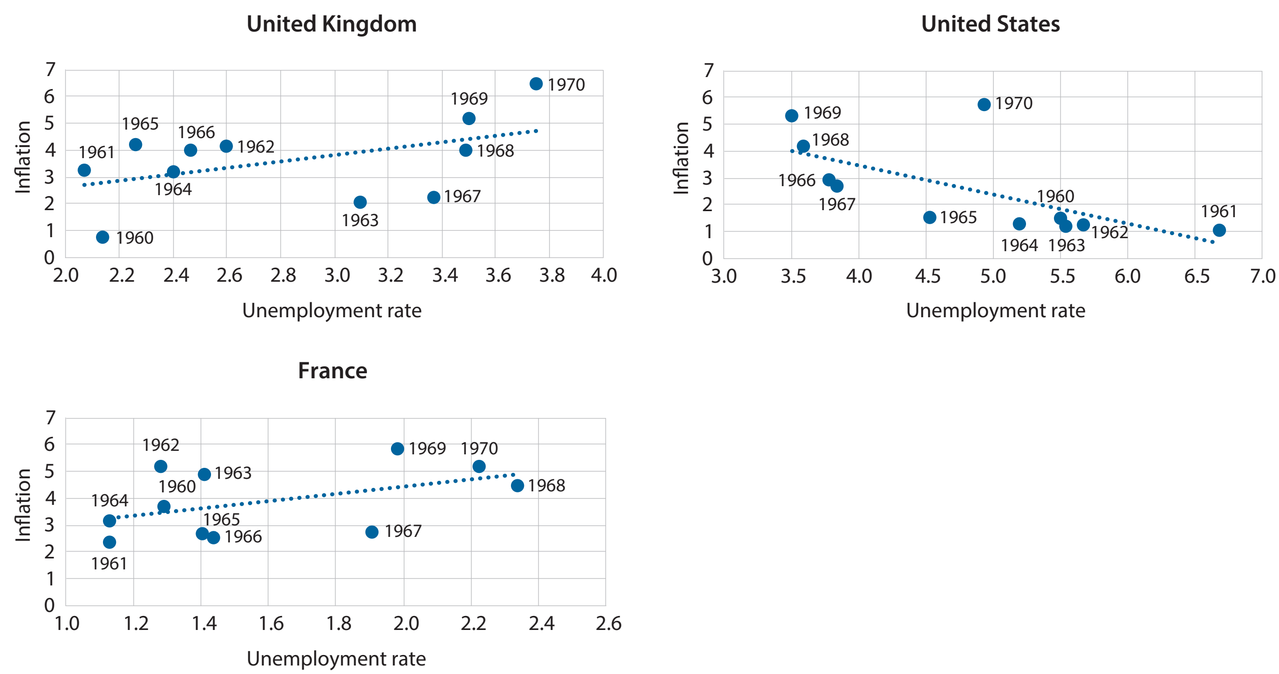
The Phillips Curve for the United States and United Kingdom, 1900-40



SOURCE: FRED®, Federal Reserve Bank of St. Louis; Annual Estimates of Unemployment in the United States, National Bureau of Economic Research; and Monthly Labor Review, Census Bureau.

Figure 2

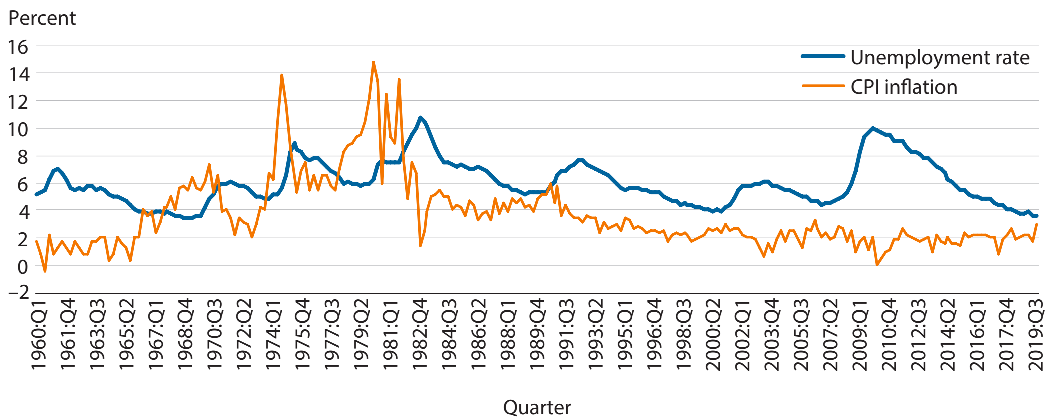
The Phillips Curve in the 1960s for Select Countries



SOURCE: FRED®, Federal Reserve Bank of St. Louis.

Figure 3

Evolution of Inflation and Unemployment Rate in the United States, 1955-2019



SOURCE: FRED®, Federal Reserve Bank of St. Louis and authors' calculations.

Phillips curve, to argue that successful monetary policy is responsible for flattening the slope of the Phillips curve by anchoring inflation expectations (Williams, 2006; Bernanke, 2007; and Mishkin, 2007). Several articles also support the argument that even in a purely static setting without expectations, the structural relationship between unemployment and inflation can be masked by the conduct of monetary policy (Bullard, 2018; Roberts, 2006; and Krogh, 2015).

3 EMPIRICAL PHILLIPS CURVE FOR THE UNITED STATES

We begin by looking at the evolution of U.S. inflation and the unemployment rate over the past 60 years. As we can see from Figure 3, the times when inflation falls below trend coincide with the times with higher unemployment. This is suggestive of a negative correlation between the series. In the 1960s, for instance, there seems to be a clear negative correlation, as well as in the 1980s. We next use some scatterplots to understand the correlation between these variables better.

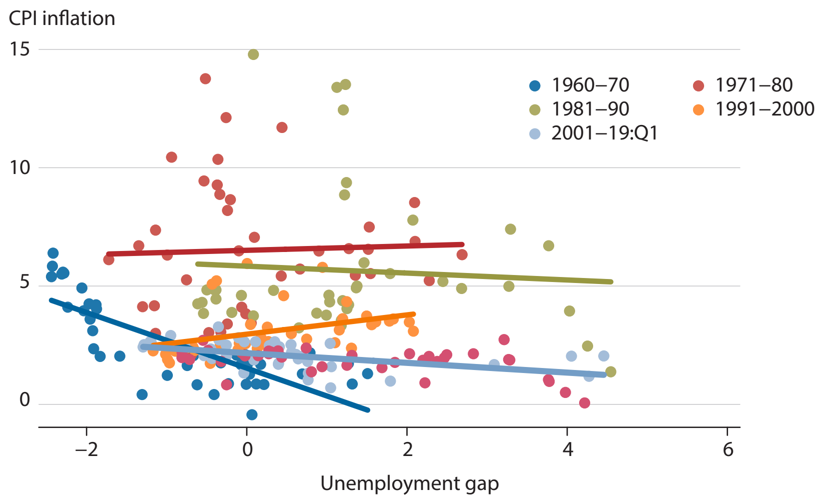
Figure 4 shows the behavior of the Phillips curve in the United States. Each dot in the graph represents a quarter. From the graph, it is clear that the slope of the Phillips curve has not been stable over the past 60 years. In the 1960s, there was a sizeable negative correlation between inflation and the unemployment gap, which has since flattened. We observe that over the past two decades, the negative relationship has returned somewhat, but the slope is not as steep as it was in the 1960s.

We formally estimate the empirical Phillips curve using aggregate data. The equation that we estimate is

$$\pi_t = \alpha + \beta_1 \left(U_t - U_t^* \right) + \beta_2 E_t \pi_{t+1} + \Sigma_{k=1}^{3} \beta_{2+k} \pi_{t-k} + \varepsilon_t ,$$

Figure 4

The U.S. Phillips Curve



NOTE: The unemployment gap is defined as the difference between the unemployment rate and the CBO estimate of natural rate of unemployment.

SOURCE: BLS, CBO, and authors' calculations.

Table 1

Phillips Curve Using Aggregate U.S. Data (1990:Q1-2018:Q1)

Inflation	Bivariate	With lags	New Keynesian	Hybrid
Unemployment gap	−0.222*** (−3.96)	−0.0171 (−0.41)	−0.159*** (−3.95)	−0.112* (−2.46)
Inflation expectations			1.006*** (12.37)	0.634*** (3.91)
Inflation lags				
First lag		0.424*** (4.32)		0.186 (1.65)
Second lag		0.461*** (4.28)		0.137 (1.36)
Third lag		0.0884 (0.86)		−0.000435 (−0.00)
Constant	2.587*** (26.32)		−0.339 (−1.51)	−0.111 (−0.47)
Observations	118	118	107	107
Adjusted R^2	0.085	0.957	0.517	0.534

NOTE: The *t*-statistics are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

SOURCE: BLS, CBO, the Survey of Professional Forecasters, and authors' calculations.

where π_t is the inflation rate and $(U_t - U_t^*)$ is the unemployment gap. We compute inflation as the annualized change in quarterly core consumer price index (CPI) inflation (obtained from the Bureau of Labor Statistics [BLS]); the unemployment gap is defined as the difference between the unemployment rate (from the BLS) and the natural rate of unemployment (from the Congressional Budget Office [CBO]). Inflation expectations are obtained from the Survey of Professional Forecasters.[1] We estimated four specifications of this Phillips curve: a bivariate model, in which we simply regress inflation on the unemployment gap ($\beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$); a model with lags, in which we regress inflation on the unemployment gap and three lags of inflation ($\beta_2 = 0$); a model with inflation expectations, in which we regress inflation on the unemployment gap and expectations ($\beta_3 = \beta_4 = \beta_5 = 0$); and a hybrid model where both past inflation and inflation expectations can influence the inflation rate.

Across all specifications, the regression coefficient for the unemployment gap is negative, which agrees with our previous discussion on the Phillips curve. The estimates in Table 1 show that the steepest slope is with the bivariate specification (–0.2), while the one with lags yields the flattest slope (–0.02). Also worth noting is the fact that inflation expectations are significant and positive in the third regression, indicating that if inflation is expected to increase in the future, current inflation responds by increasing as well.

4 THE BASIC NEW KEYNESIAN FRAMEWORK

To better understand the problems and limitations of a simple empirical estimation of the slope of the Phillips curve, we now study the issue using a macroeconomic model: the New Keynesian model. We first describe the canonical New Keynesian model and then discuss its equilibrium conditions. We follow closely Clarida, Gali, and Gertler (1999), as well as Chapter 3 of Gali (2008).

Within the New Keynesian macro-model framework, monetary policy affects the real economy in the short run. The basic model consists of identical households that make decisions about labor supply, consumption, and savings. Households are risk averse and dislike fluctuations in consumption. Households can borrow or save in one-period bonds that pay an interest rate. These bonds are in zero net supply. Firms produce a range of differentiated goods using only labor with a constant returns-to-scale production function. Firms set prices for their goods, taking into account that they face competition from close substitutes of their goods and that they are unable to change prices every period and able to do so only after some random period. In the basic model, we abstract from investment and capital and from government consumption and interactions with the rest of the world. The total production cannot be stored, and total consumption has to equal total production. We also use a version of the model that abstracts from money holdings, sometimes referred to as the cashless limit economy. Since all households are identical, in equilibrium the interest rate must be such that households are indifferent between saving and consuming their income.[2]

The dynamic conditions that characterize equilibrium in this model can be approximated by two equations: a dynamic IS curve that relates the output—in deviations from trend—to the real interest rate, and a Phillips curve that relates inflation positively to the output. It is important to note this dynamic aggregate behavior evolves from optimization by firms and households. Moreover, these equations approximate the dynamics around a stationary equilibrium.

For any variable z, let $\hat{z}$ denote the deviation of the variable from its steady state or long-run value. The New Keynesian Phillips curve is given by

$$\hat{\pi}_t = \beta E_t\left[\hat{\pi}_{t+1}\right] + \kappa \hat{y}_t + \varepsilon_t. \tag{1}$$

Here, π_t refers to the aggregate inflation at time t, $\hat{y}_t$ refers to the output in deviations from trend or potential at time t, and ε_t is a cost-push shock that follows an AR(1) process ($\varepsilon_t = \rho\varepsilon_{t-1} + e_t$).

This equation reflects the aggregate relationship between the inflation rate—in deviations from a long-run value—and output, and it tells us that an increase in output is associated positively with a positive change in inflation. This equation evolves from the individual firm's problem: Firms are monopolistically competitive, and each firm chooses its price to maximize profits subject to constraints on the frequency of future price adjustments. The aggregate Phillips curve is simply capturing the aggregation of individual pricing decisions by a log-linear approximation. Another way of interpreting this equation is by thinking in terms of marginal costs. Excess demand (output above the potential) is associated with marginal costs above average, and the reverse is true in the case of excess supply. Thus, firms set nominal prices based on their expectations of future marginal costs, as well as current marginal costs. The cost-push shock is meant to capture other forces that may affect inflation not already captured in the model.

The next ingredient of the model is the dynamic IS curve, which inversely relates the aggregate output (or consumption since they are the same in this model) to the real interest rate, and tells us that aggregate output evolves as a function of the expected future output, the interest rate, the expected inflation level, and demand shocks:

$$\hat{y}_t = E_t\left[\hat{y}_{t+1}\right] - \left(i_t - E_t\left[\hat{\pi}_{t+1}\right] - z_t\right).$$

Here, z_t is a demand shock that follows an exogenous AR(1) process with persistence ρ_z (i.e., $z_t = \rho_z z_{t-1} + \gamma_t$). This equation is a result of log linearizing the consumption Euler equation that arises from optimizing the household's utility. The term $i_t - E_t[\hat{\pi}_{t+1}]$ is the (expected) real interest rate since i_t is the nominal rate. The mechanism by which the real interest rate affects output is somewhat intuitive: Higher interest rates imply higher borrowing costs or a higher return to savings, which lead to lower consumption today and thus lower current output. This is also known as the intertemporal substitution of consumption. The reason that expected future output affects current output is that individuals prefer a smooth path of consumption. If individuals expect higher income and consumption in the next period, they prefer to consume more today, thus resulting in an increase in output today. Overall, output demand depends negatively on the real interest rate and positively on expected future output. One possible interpretation for the demand shock z_t is a change in households' preferences for consuming today versus tomorrow, like a sudden increase in impatience.

This simple model does not have a direct link to unemployment. We add an exogenous relationship between the output and the unemployment rate, or, more precisely, the unemployment rate gap, which is best known as Okun's law:

$$\hat{u}_t = \alpha \hat{y}_t.$$

Finally, as is usual in most New Keynesian models, we add a monetary authority with the objective to control inflation and promote high levels of employment. The monetary authority sets the nominal interest rate to achieve these objectives. In particular, we assume that the short-term nominal interest rate is set according to a Taylor rule of the form

$$i_t = r^* + \pi^* + \phi_\pi \ \hat{\pi}_t + \phi_y \ \hat{y}_t,$$

where r^* is the long-run, or natural, real interest rate, and π^* is a long-run value of inflation, both of which we assume are exogenous. We assume that the coefficients ϕ_π and ϕ_y are positive. This policy function implies that the monetary authority will increase the nominal interest rate if inflation is above its long-run value—that is, when $\hat{\pi}_t$ is positive—or when output is above its long-run value—that is, when $\hat{y}_t$ is positive. A solution to this model includes decision rules for inflation, output, interest rate, and unemployment, as a function of the two exogenous shocks, demand shocks, and cost-push shocks. Assuming that the Taylor principle holds and that the parameters of the model are within normal ranges, we can solve for the equilibrium inflation and output by plugging this optimal policy rule into the Phillips curve. We then find that both the equilibrium inflation and output gap are functions of the demand shock and cost-push shock:

$$\hat{y}_t = c_{y\varepsilon}\varepsilon_t + c_{yz}z_t,$$

$$\hat{i}_t = c_{i\varepsilon}\varepsilon_t + c_{iz}z_t,$$

$$\hat{\pi}_t = c_{\pi\varepsilon}\varepsilon_t + c_{\pi z}z_t, \text{ and}$$

$$\hat{u}_t = c_{u\varepsilon}\varepsilon_t + c_{uz}z_t,$$

where coefficients in the equations are constants that depend on model parameters. In the analysis that follows, we focus on the last two equations that say that the equilibrium values of inflation and unemployment at time t depend on the realization of shocks. In particular, it can be shown that both of the coefficients in the inflation decision rule are positive; thus, inflation increases with positive demand shocks or positive cost-push shocks. On the other hand, unemployment decreases with positive demand shocks but increases with cost-push shocks, as these last shocks will induce a contraction in economic activity.

We now use this model as a data-generating process to simulate shocks to the economy. For this, we need to specify values for the parameters of the model. We use the same calibration as Gali (2008).[3] We can use the New Keynesian Phillips curve in equation (1), together with the equation for Okun's law, to derive a relationship between inflation and unemployment, which we can estimate. In particular, we have

$$(2) \qquad \hat{\pi}_t = \beta E_t\left[\hat{\pi}_{t+1}\right] + \delta\, \hat{u}_t + \varepsilon_t,$$

where the coefficient δ is a combination of parameters κ and α. A direct estimation of equation (2) is problematic. As the decision rules clearly show, a simple regression of inflation on the unemployment gap would give us biased estimates because the unemployment gap is correlated with the cost-push shock ε_t. In a similar way, the expected value of inflation in the next period would also be correlated with today's cost-push shock if the shocks were persistent.

McLeay and Tenreyro (2019) discuss several assumptions and estimation methods under which estimates would be correct. First, in an economy without cost-push shocks and in which expected inflation and unemployment are observed, a regression would recover the slope δ. Alternatively, if there are good instruments for both the unemployment rate and the expected inflation that are uncorrelated to the cost-push shock, then an instrumental variables regression would be valid. Finally, they propose a different way to tackle the problem, which we discuss later.

We now use the model to understand possible shortcomings in the empirical estimation of Phillips curves. The exercises that we describe next use the decision rules for inflation and unemployment together with simulated values for the demand and cost-push shocks. These give a simulated time series for the two variables in the Phillips curve. Figure 5 shows the results of simulating various combinations of shocks. To simplify the analysis and make our point, we assume that the shocks are not persistent. In other words, we assume that $\rho = \rho_z = 0$, which would then imply that $E_t[\pi_{t+1}]$ is not affected by the current realization of the shocks.[4]

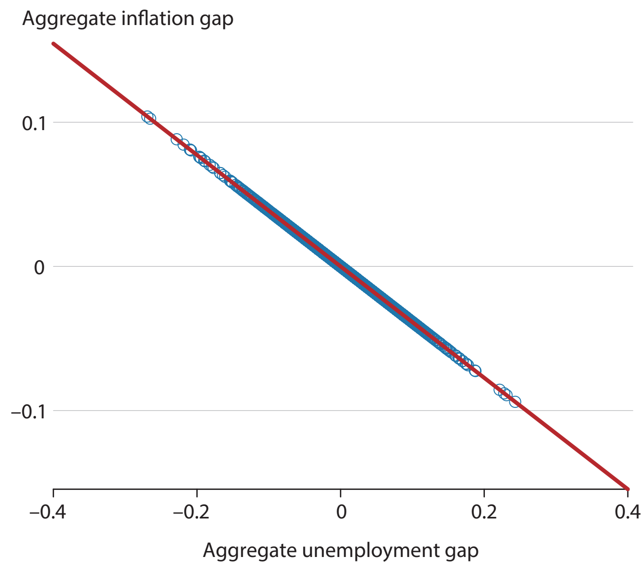
4.1 Case I: Only Demand Shocks

We now start our analysis by focusing on the effects of different forces on the slope of the Phillips curve, one by one. In this case, we allow for only demand shocks. As Figure 5 shows, all the different equilibrium values of inflation and unemployment reflect a typical negative relationship between inflation and unemployment.

This is what we expect to see given the decision rules obtained before and that only demand shocks are allowed. When a demand shock hits, output increases, lowering unemployment but also increasing inflation. The monetary authority counteracts positive demand shocks by increasing the interest rate, which pushes both output and inflation to an otherwise lower level. In this case, the monetary authority does not face a trade-off between output and inflation. Note, however, that the slope of the empirical Phillips curve in Figure 5, in general, will not be equal to the slope of the structural Phillips curve—that is, the coefficient δ in equation (2). The reason is that the actions of the monetary authority affect the movements in inflation and the unemployment rate, which in turn affect the empirical slope, as we highlight in Case III.

Figure 5

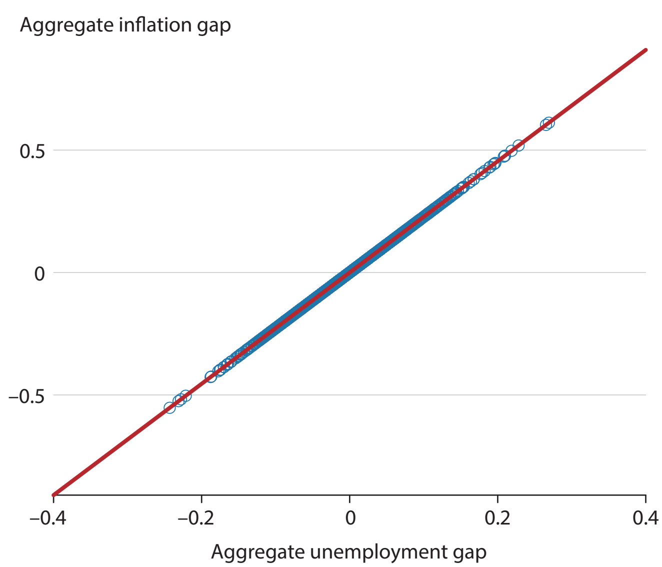
Simulated Inflation and Unemployment with Only Demand Shocks



SOURCE: Authors' calculations.

Figure 6

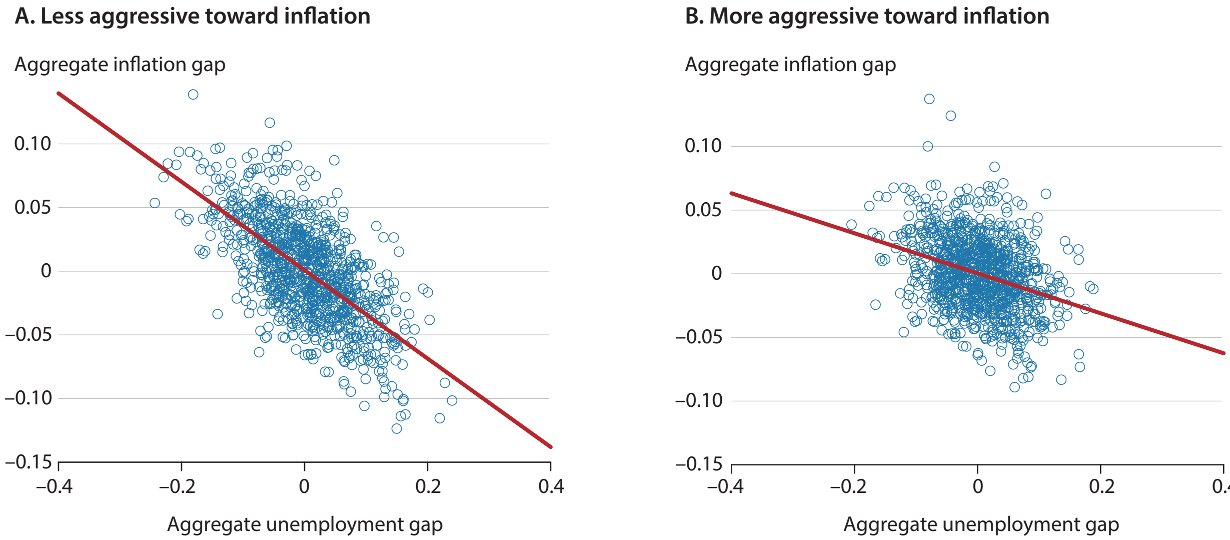
Simulated Inflation and Unemployment with Only Cost-Push Shocks



SOURCE: Authors' calculations.

Figure 7

Simulated Inflation and Unemployment with Different Policy Parameters



SOURCE: Authors' calculations.

4.2 Case II: Only Cost-Push Shocks

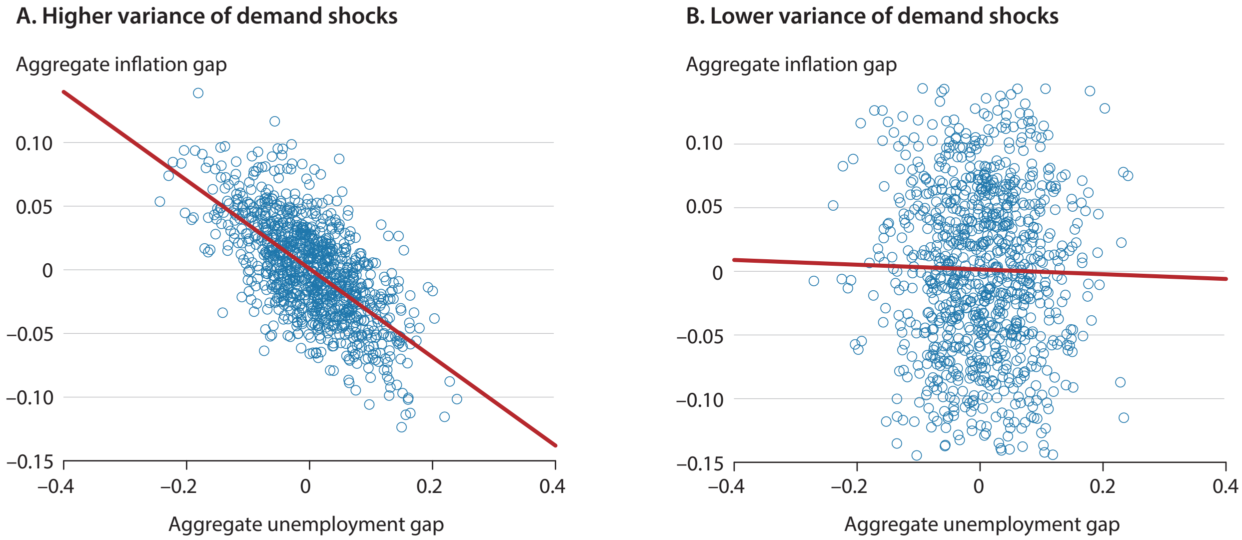
We next set all demand shocks equal to zero in our simulations and allow for only cost-push shocks. Figure 6 shows the results of our simulations for this case. In the decision rules, the coefficient that multiplies the cost-push shock is positive for both variables—unemployment and inflation. Thus, with a positive cost-push shock, both inflation and unemployment increase and we observe a positive relationship between inflation and the unemployment gap in the graph. In the background, it is the optimal decisions of agents and the monetary authority that drive this behavior. In particular, the monetary authority is aggressive at fighting inflation resulting from the cost-push shock, increasing the interest rate, thus lowering total consumption and production and increasing unemployment.

4.3 Case III: Changes in the Policy Rule Over Time

In a recent discussion, Bullard (2018) studies how the empirical Phillips curve would look in a simple New Keynesian model if the monetary authority reacts more aggressively toward increases in inflation. Figure 7 shows this comparison. Panel A presents simulations using a value of $\phi_\pi = 1.5$, and Panel B uses a value of $\phi_\pi = 5.5$, which implies a more aggressive stance toward increases in inflation. As the figure shows, with a more aggressive policy, the empirical slope of the Phillips curve is flatter. The main takeaway of this exercise is that if the monetary authority becomes more aggressive toward inflation over time, then, all else equal, the empirical Phillips curve will be flatter. However, this flattening of the empirical Phillips curve occurs even if the structural Phillips curve of equation (1) did not change, as we are keeping δ constant.

Figure 8

Simulated Inflation and Unemployment with Different Variance of Cost-Push Shocks



SOURCE: Authors' calculations.

This exercise illustrates an important pitfall in empirical analysis using reduced-form estimations of the Phillips curve: The actions of the monetary authority affect estimates of the slope of the Phillips curve. Even if the slope of the structural Phillips curve remains constant (parameter δ), changes in monetary policy over time affect the empirical slope of the curve, and a more aggressive stance will make the empirical curve flatter.

4.4 Case IV: Changes in the Variance of Shocks

As Cases I and II show, the slope of the empirical Phillips curve will be affected by the type of shock that impacts the economy. We next simulate both shocks and compute the corresponding time series behavior for inflation and unemployment according to the decision rules. Figure 8 presents the results for two different sets of parameters. In particular, Panel A shows the simulations for an economy with a relatively higher variance of the cost-push shock, while Panel B shows the same simulations but for a lower variance of these shocks. As is clear from the figure, an economy with larger cost-push shocks will display a flatter (or even an upward-sloping) empirical Phillips curve. Once again, it is worth stressing that the flattening of the empirical Phillips curve would occur even if the structural Phillips curve of equation (2) does not change and δ remains constant, which also shows a potential pitfall in empirical analysis using reduced-form estimations of the Phillips curve. To the extent that empirical estimates cannot control properly or instrument correctly for cost-push shocks, differences in the variances of the shocks will be reflected in the estimates of the slope.

Case III highlights that the empirical Phillips curve may flatten over time if the monetary authority adopts a more aggressive stance toward inflation. Case IV, on the other hand, highlights that the flattening could be due to a decrease in the importance (size) of demand shocks relative to cost-push shocks.

5 REGIONAL PHILLIPS CURVES

As we mentioned before, McLeay and Tenreyro (2019) discuss several avenues to properly recover the slope of the Phillips curve. The first method they suggest is to control for cost-push shocks, to the extent that these are observable. In essence, this would allow us to replicate Case II from the previous section by removing the effect of cost-push shocks from both inflation and the unemployment rate, such that any remaining variation in the unemployment gap and inflation gap must be due to movements in aggregate demand. There are articles, such as Roberts (1995), which use oil prices to control for cost-push shocks. This approach, while reasonable, is difficult to implement in practice since there is a large number of potential cost-push shocks that would need to be controlled for, which in many cases may not be observable.

Another method is to use an instrumental variable estimation for the unemployment gap. The instrument should be correlated with unemployment but uncorrelated with cost-push shocks. This is also hard to do in practice because it is difficult to find good instruments—that is, a macroeconomic variable that is truly exogenous to cost-push shocks but correlated with the unemployment rate and inflation expectations. Moreover, strong temporal dependence or persistence in the shocks may also affect the validity of the instruments.[5]

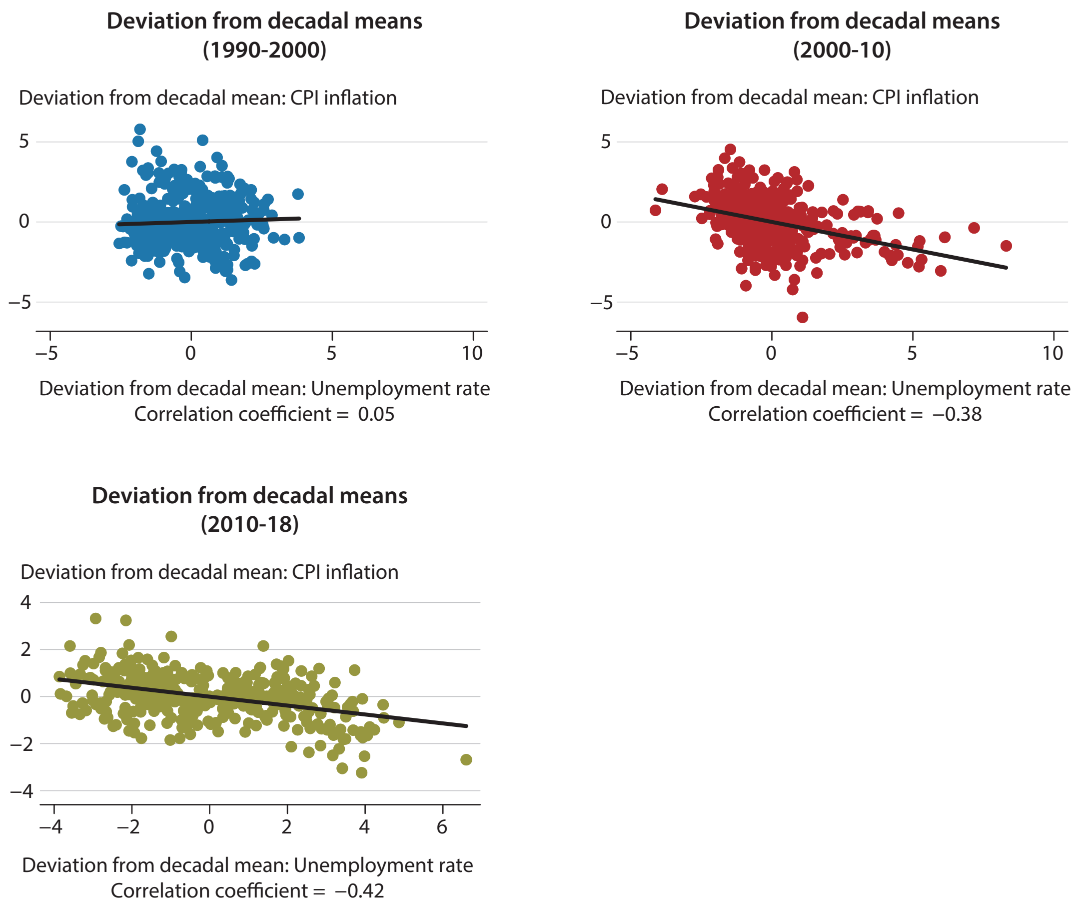
Finally, the third method they mention, and that we analyze in this article, is to exploit the cross-sectional heterogeneity in regional data. Fitzerald and Nicolini (2014) also implement this approach in their article. As highlighted in Case III, the actions of the monetary authority affect the movements in unemployment and inflation and thus affect the estimates of the slope of the Phillips curve, even if properly instrumenting for cost-push shocks. However, in a large country like the United States, comprising several states and cities, the monetary authority only reacts to fluctuations in aggregate inflation and unemployment, and monetary policy is independent of local conditions. Thus, the main idea behind a regional approach to estimating the Phillips curve is to use fluctuations in regional economic conditions, in deviations from the aggregate, which the monetary authority will not influence by its actions. In this way, by properly instrumenting or controlling for cost-push shocks, we should be able to recover the structural relationship between inflation and the unemployment rate as we are also controlling for monetary policy. We first discuss empirical evidence and then use a model to analyze this approach.

5.1 Empirical Analysis

To explore the empirical implications of this regional approach to the Phillips curve estimation, we use semi-annual CPI inflation and unemployment data from the BLS for 20 metropolitan statistical areas (MSAs). We divide our sample into three periods: 1990-2000, 2000-10,

Figure 9

Empirical Relation between the Inflation Gap and Unemployment Gap for Different MSAs



SOURCE: BLS and authors' calculations.

and 2010-19. For each decade, we compute the mean inflation and unemployment for each MSA and use this as a proxy for the steady-state level of unemployment and inflation for each MSA. In Figure 9 we use simple scatterplots of the deviation of inflation from the decadal mean versus the deviation of unemployment from the decadal mean for each MSA. We interpret these as the unemployment and inflation gaps for each MSA.

As can be seen from Figure 9, the slope of the scatterplots varies over time. The curve was positively sloped in the 1990s, which contradicts our understanding of the theoretical relationship between these variables.

Formally, we estimate

$$\pi_{it} = \alpha_i + \beta_1 E_t \pi_{it+1} + \beta_2 U_{it} + \gamma_t + \varepsilon_{it},$$

Table 2

Estimates of the Empirical Phillips Curve Using MSA-Level Data, 1990-2018

Inflation	Pooled	Metro area FE only	Year FE only	Metro area and year FE
Unemployment rate	−0.141*** (−7.20)	−0.168*** (−6.57)	−0.215*** (−5.81)	−0.406*** (−5.91)
Inflation lag				
First lag	0.390*** (8.00)	0.367*** (6.97)	0.127* (2.58)	0.0570 (1.18)
Within R^2	0.252	0.254	0.435	0.446
Observations	1,191	1,191	1,191	1,191
MSA FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Seasonal dummies	Yes	Yes	Yes	Yes

NOTE: FE, fixed effects. The *t*-statistics are in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
SOURCE: BLS.

where subindex i denotes an MSA and subindex t denotes the time period. As before, π is the inflation rate and U is the unemployment rate. The regression includes both time fixed effects and MSA fixed effects. Including MSA fixed effects allows us to remove the effect of the state-specific mean unemployment rate, which can be seen as a proxy for the equilibrium rate of unemployment for each MSA. The time fixed effects help control for time-varying changes in the aggregate equilibrium unemployment rate and help us overcome the bias caused by the correlation between the regional unemployment rate and aggregate unemployment rate.[6] In Table 2, we present results from estimating the above regression, with a combination of fixed effects.

Column (2) shows the results of regressing inflation on unemployment without controlling for time fixed effects or seasonality. However, once we begin to control for these factors, we observe that the coefficient increases in magnitude; from Column (5), we observe that the coefficient is highly negative and significant.

5.2 A Simple New Keynesian Model of a Monetary Union

To understand the factors that influence the empirical estimates of the slope of the Phillips curve using regional data, we turn to a simple model of a monetary union with a common monetary authority, specified by Gali and Monacelli (2008). With this model, we proceed in a similar fashion as before, finding decision rules and then simulating shocks and time series for inflation and unemployment, in this case at both the aggregate and regional levels.

Gali and Monacelli (2008) model the currency union as a continuum of small (atomistic) open economies that are subject to imperfectly correlated shocks. Each of these small economies, which we call regions, share identical preferences, technology, and market structures.

Let P_t^j represent region j's price index for goods produced locally, and let P_t^f be the price index for goods purchased (imported) from other regions. In this way, the CPI for a region is the geometric weighted average of domestic and imported price indexes, with weights given by consumption shares $P_{c,t}^j = (P_t^j)^{1-\alpha}(P_t^*)^{\alpha}$. For the monetary union as a whole, the price index is the aggregator across all regions.

The bilateral terms of trade between regions j and k can be defined as $S_{f,t}^j = \frac{P_t^f}{P_t^j}$, which is the price of region k's goods in terms of region j's goods. Then, the CPI inflation rate of the region ($\pi_{c,t}^j$) can be approximated as follows:

$$\pi_{c,t}^j = \pi_t^j + \alpha \Delta s_t^j.$$

The model assumes complete markets and perfect risk-sharing across all regions, together with a zero net savings for all regions. This means that regional consumption and consumption for the whole country can be approximated as

$$c_t^j = c_t^* + (1-\alpha)\, s_t^j,$$

where c_t^* is the (log) aggregate consumption for the whole country.

A log-linear approximation of the model around a symmetric steady state for all regions leads to similar dynamic equilibrium conditions as before. For region j, the Phillips curve is

$$\pi_t^j = \beta\ E_t\left[\pi_{t+1}^j\right] + \kappa\ \hat{y}_t^j + \varepsilon_t^j.$$

In addition, in each region an IS equation holds:

$$c_t^j = E\left[c_{t+1}^j\right] - \left(i_t^* - \pi_{c,t}^j - z_t^j\right),$$

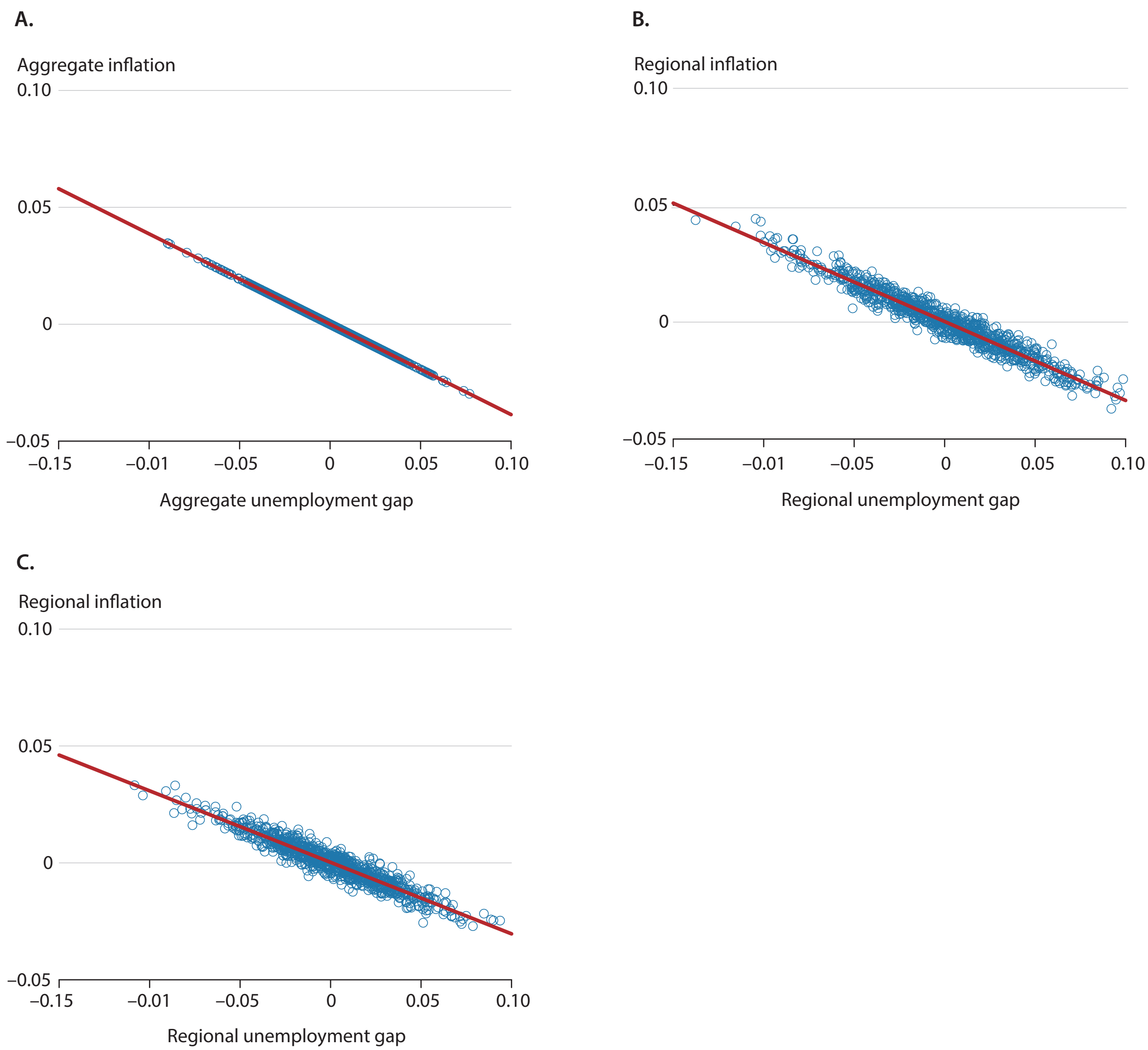
where i_t^* is the interest rate for the whole country. Aggregating these expressions across regions can form a Phillips curve and an IS curve for the whole country, which leads to the same expressions as in Section 4. We further assume that an expression like Okun's law holds in every region and that the monetary authority follows the same Taylor rule as in Section 4, reacting only to fluctuations in aggregate inflation and the aggregate output gap. As before, we use this model to simulate shocks to the economy.[7]

5.3 Case V: Only Demand Shocks—Regional Analysis

As in Case I, here we allow for only demand shocks. We assume a few possible scenarios. First, we allow for only demand shocks, both at the aggregate level and for individual regions. Second, we allow for only demand shocks at the regional level, but no aggregate shocks of any kind. In this last case, aggregate inflation and unemployment are constant, and the monetary authority will not change the interest rate. Figure 10 presents the results. Panel A shows the effect of aggregate demand shocks at the aggregate level—that is, for aggregate inflation and the aggregate unemployment rate. Panel B shows the effect of aggregate and regional demand shocks at the regional level. That is, each dot in the figure shows the inflation rate and the

Figure 10

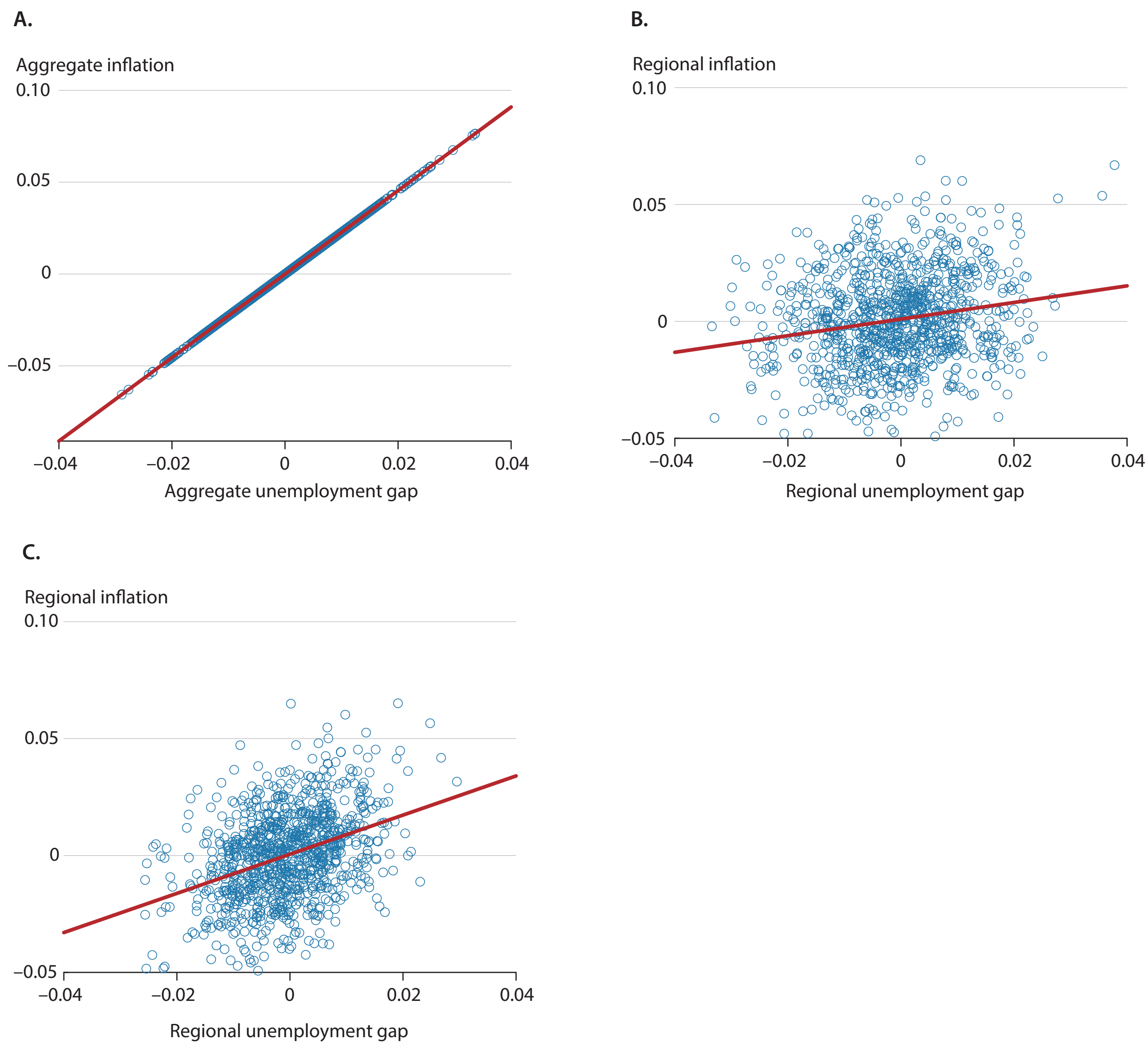
Effects of Regional and Aggregate Demand Shocks



SOURCE: Authors' calculations.

Figure 11

Effects of Regional and Aggregate Cost-Push Shocks



SOURCE: Authors' calculations.

unemployment rate in one individual region. Finally, Panel C shows the effect of only regional demand shocks at the regional level. In this case, the aggregate inflation and unemployment rate are constant at their steady-state value.

Clearly, for aggregate inflation and unemployment displayed in Panel A, allowing for only demand shocks shows a similar picture as Case I (shown in Figure 5). Adding regional shocks on top of this (Panel B) introduces some noise around the empirical relationship, but the main pattern still holds. More interesting, Panel C shows the relationship in the absence of aggregate shocks. In this case, we can see that regional demand shocks can help identify a negative empirical slope between inflation and unemployment at the regional level.

5.4 Case VI: Only Cost-Push Shocks—Regional Analysis

Similar to the previous case, we now analyze the effects of allowing for only cost-push shocks. Figure 11 shows the results. In Panel A, and similar to Figure 6, cost-push shocks generate a positive-sloping empirical relationship between inflation and unemployment for the aggregate economy. While the monetary authority intervenes to stabilize the economy, it faces the trade-off between lower inflation at the expense of a larger contraction and higher unemployment. At the regional level, the effects of both aggregate and regional cost-push shocks, shown in Panel B, generate a more dispersed pattern. Yet, the empirical slope between unemployment and inflation is still positive when using only regional data. Finally, Panel C shows the results for an economy with only regional cost-push shocks. In this case, aggregate variables do not move and there is no intervention by the monetary authority; the positive slope is more pronounced.

This exercise leads to an important conclusion. The empirical analysis using regional data must also control or instrument for regional cost-push shocks. Otherwise, estimates of the Phillips curve will be affected.

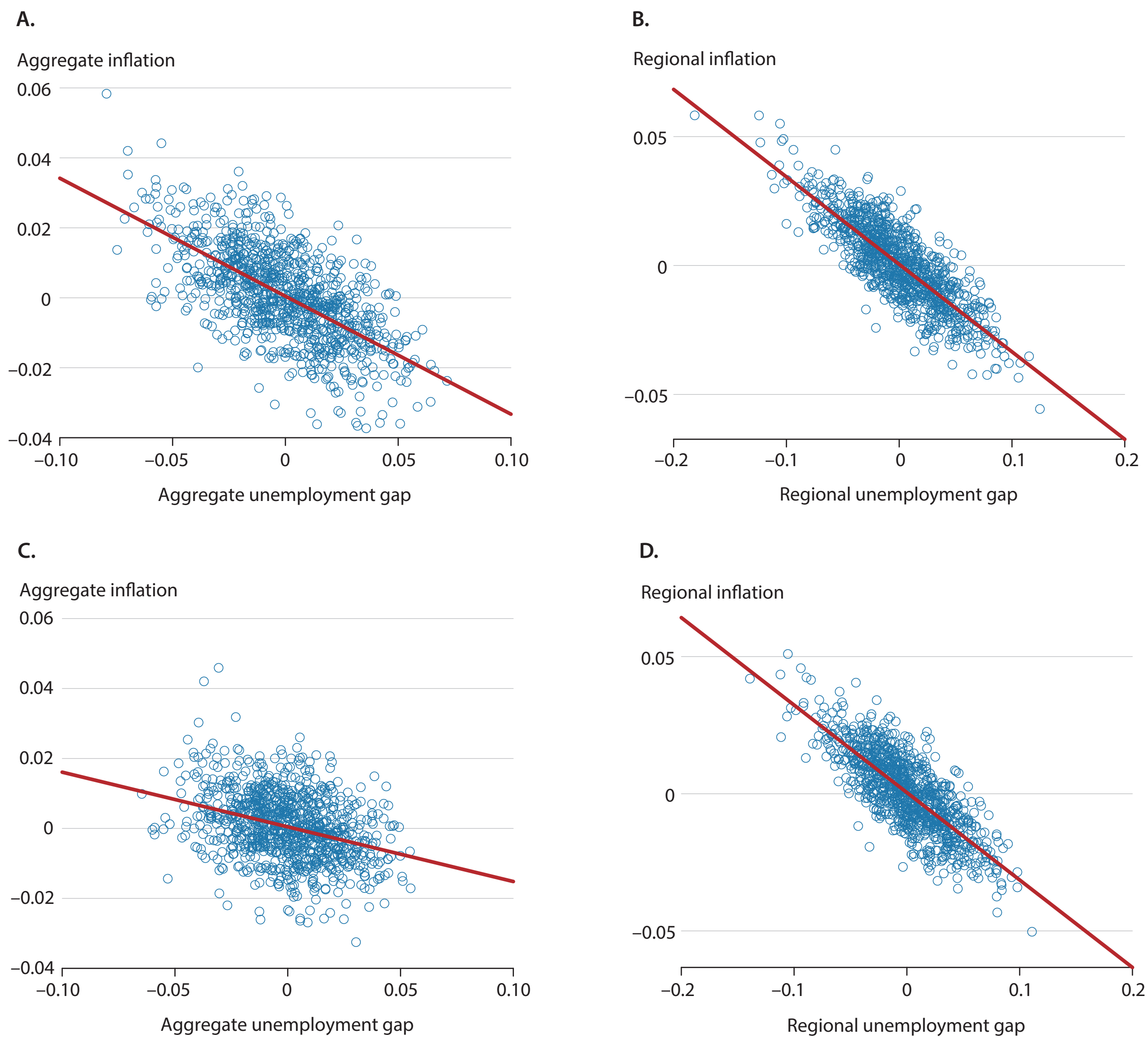
5.5 Case VII: Changes in the Policy Rule Over Time—Regional Analysis

Similar to our methods in the simple model, here we ask if changes in the monetary authority's preferences for how aggressively to fight inflation affect the empirical slope of the Phillips curve estimated using regional data. For this we use an economy where all shocks are active (aggregate and regional; demand and cost-push), and we compare whether and how the empirical slope of the Phillips curve estimated using regional data changes. We show this comparison in Figure 12, where we use simulated data from the model.

Panels A and C show the relation between unemployment and inflation at the aggregate level for two economies with different degrees of monetary policy aggressiveness, with Panel C showing the results for an economy with a more aggressive stance. These graphs are similar to those in Case III analyzed above, where the empirical slope of the Phillips curve was affected by changes in the monetary authority's preferences. Panels B and D show the relationship between unemployment and inflation using regional data, where Panel D shows the case of a monetary authority more aggressive toward inflation. In this case, the changes in the policy rule do not translate into noticeable changes in the slope of the empirical Phillips curve at the regional level. In other words, while the slope estimated with regional data does change with

Figure 12

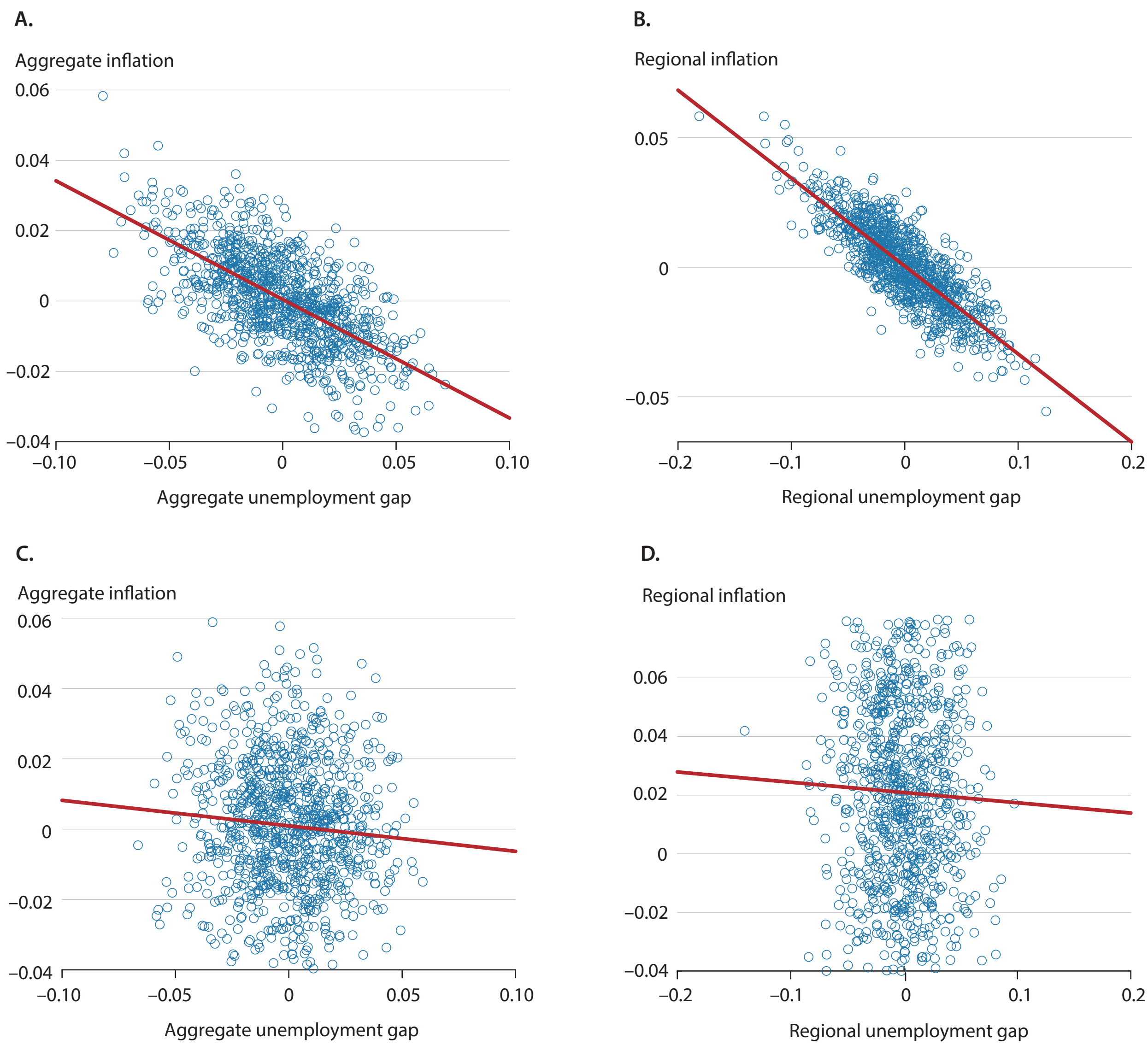
Effects of Changes in the Policy Rule at Aggregate and Regional Levels



SOURCE: Authors' calculations.

Figure 13

Effects of Changes in the Variance of Shocks at Aggregate and Regional Levels



SOURCE: Authors' calculations.

changes in monetary policy, the changes are very small. This result lends some support to the use of regional data for estimating the slope of the curve, as it is almost invariant to changes in policy.

5.6 Case VIII: Changes in the Variance of Shocks—Regional Analysis

Finally, we analyze the effect of changes in the relative variance of demand and supply shocks to answer the following question: Can regional data help estimate the slope of the empirical Phillips curve if the relative variance of shocks changes? We proceed in a similar way as in Case IV presented earlier and increase the variance of cost-push shocks relative to the variance of demand shocks, at both the aggregate and regional levels. Figure 13 shows the results of this exercise. Panels A and B display the baseline case for aggregate and regional variables, respectively; and Panels C and D have the same information but for an economy with a larger variance of cost-push shocks. Comparing the top and bottom panels makes it clear that an increase in the relative variance of cost-push shocks will affect the estimated slope of the empirical Phillips curve, whether estimated using aggregate data or regional data.

This highlights an important limitation with the use of regional data to estimate Phillips curves. The empirical slope of the curve does not need to capture the structural slope in equation (2). Moreover, the empirical slope may flatten over time even if the structural slope did not change.

6 CONCLUSION

We use a simple New Keynesian framework to illustrate the main problems in estimating the slope of the structural Phillips curve. Some of these problems arise due to the actions of monetary policy, as it affects economic activity to fight inflation, which leads to biases in simple estimations. Recent proposals have favored estimations using regional or city data in an effort to overcome these issues, as monetary policy will react to only aggregate economic conditions and not regional. We use a simple model of a monetary union with a continuum of economies and a common monetary policy authority that reacts to aggregate conditions. When we use this model as a data-generating process, we find that the main drawbacks of the aggregate model are still present in a cross-section of many regions in a monetary union. The relative importance of the demand and supply shocks will largely determine the empirical relation between unemployment and prices in both the aggregate and the cross-section of regions. Our analysis shows potential pitfalls in estimating the slope of the Phillips curve, even if using regional data. ■

NOTES

[1] We follow McLeay and Tenreyro (2019) and use 10-year-ahead mean forecasts for CPI inflation as our measure of inflation expectations. In their words, "we use five to ten year ahead inflation expectations, as suggested by Bernanke (2007) and Yellen (2015) as having a stronger empirical fit with the data" (p. 32).

[2] For further details of the model, see Gali (2008).

[3] The specific values we use in our quantitative analysis are not important. Our goal is to make a qualitative point, not a quantitative one.

[4] Note, however, that this assumption is not without loss of generality, since, as we mentioned before, with persistent shocks future expected inflation is correlated with the current realization of the shocks.

[5] And, as said before, one must also instrument for expected inflation.

[6] Note that in the regional Phillips curve case, we abstract from inflation expectations. The reason is that we do not have data on inflation expectations at the regional level, since the Survey of Professional Forecasters contains information on inflation for the whole U.S. economy.

[7] Note that we assume regions within the country are small and do not affect aggregate variables. This assumption is not without loss of generality. In fact, this assumption gives the best chance of success to the hypothesis that using regional data would recover the slope of the Phillips curve. Our point is that, even in this case, the assumption may fail in some circumstances.

REFERENCES

Atkeson, Andrew and Ohanian, Lee E. "Are Phillips Curves Useful for Forecasting Inflation?" Federal Reserve Bank of Minneapolis *Quarterly Review*, 2001, *25*(1), pp. 2-11; https://doi.org/10.21034/qr.2511.

Bernanke, Ben. "Inflation Expectations and Inflation Forecasting." Speech delivered at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, MA, July 19, 2007; https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm.

Blanchard, Olivier; Cerutti, Eugenio and Summers, Lawrence. "Inflation and Activity–Two Explorations and their Monetary Policy Implications." Working Paper No. 21726, National Bureau of Economic Research, 2015; https://doi.org/10.3386/w21726.

Bullard, James. "The Case of the Disappearing Phillips Curve." Speech delivered at the 2018 ECB Forum on Central Banking, Sintra, Portugal, June 19, 2018; (accompanying slides) https://www.stlouisfed.org/~/media/files/pdfs/bullard/remarks/2018/bullard_ecb_sintra_june_19_2018.pdf.

Calvo, Guillermo A. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics*, 1983, *12*(3), pp. 383-98; https://doi.org/10.1016/0304-3932(83)90060-0.

Cecchetti, Stephen G.; Feroli, Michael; Hooper, Peter; Kashyap, Anil K. and Schoenholtz, Kermit. "Deflating Inflation Expectations: The Implications of Inflation's Simple Dynamics." CEPR Discussion Paper No. DP11925, Centre for Economic Policy Research, March 2017.

Clarida, Richard; Gali, Jordi and Gertler, Mark. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature*, 1999, *37*(4), pp. 1661-707; https://doi.org/10.1257/jel.37.4.1661.

Coibion, Olivier and Gorodnichenko, Yuriy. "Is the Phillips Curve Alive and Well After All? Inflation Expectations and the Missing Disinflation." *American Economic Journal: Macroeconomics*, 2015, *7*(1), pp. 197-232; https://doi.org/10.1257/mac.20130306.

Fitzgerald, Terry J. and Nicolini, Juan Pablo. "Is There a Stable Relationship Between Unemployment and Future Inflation? Evidence from U.S. Cities." Working Paper 713, Federal Reserve Bank of Minneapolis, Research Department, May 2014; https://doi.org/10.21034/wp.713.

Galí, Jordi. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, 2008.

Galí, Jordi and Monacelli, Tommaso. "Optimal Monetary and Fiscal Policy in a Currency Union." *Journal of International Economics*, 2008, *76*(1), pp. 116-32; https://doi.org/10.1016/j.jinteco.2008.02.007.

Lucas Jr., Robert E. "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 1972, *4*(2), pp. 103-24; https://doi.org/10.1016/0022-0531(72)90142-1.

McLeay, Michael and Tenreyro, Silvana. "Optimal Inflation and the Identification of the Phillips Curve." NBER Working Paper 25892, National Bureau of Economic Research, 2019; https://doi.org/10.3386/w25892.

Mishkin, Frederic S. "Inflation Dynamics." *International Finance*, 2007, *10*(3), pp. 317-34; https://doi.org/10.1111/j.1468-2362.2007.00205.x.

Phillips, Alban W. "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957." *Economica*, 1958, *25*(100), pp. 283-99; https://doi.org/10.2307/2550759.

Roberts, John M. "New Keynesian Economics and the Phillips Curve." *Journal of Money, Credit, and Banking*, 1995, *27*(4), pp. 975-84; https://doi.org/10.2307/2077783.

Roberts, John M. "Monetary Policy and Inflation Dynamics." *International Journal of Central Banking*, September 2006, *2*(3).

Samuelson, Paul A. and Solow, Robert M. "Analytical Aspects of Anti-Inflation Policy." *American Economic Review*, 1960, *50*(2), pp. 177-94.

Stock, James H. and Watson, Mark W. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 2007, *39*, pp. 3-33; https://doi.org/10.1111/j.1538-4616.2007.00014.x.

Williams, John C. "Inflation Persistence in an Era of Well-Anchored Inflation Expectations." Federal Reserve Bank of San Francisco *Economic Letter*, Number 2006-27, October 2006; https://www.frbsf.org/economic-research/publications/economic-letter/2006/october/inflation-persistence-in-an-era-of-well-anchored-inflation-expectations/.